

P. Morfeld^{1,2} · M. Spallek^{3,4}¹ Institut und Poliklinik für Arbeitsmedizin, Umweltmedizin und Präventionsforschung der Universität zu Köln, Köln, Deutschland² Institut für Epidemiologie und Risikobewertung in der Arbeitswelt (IERA) der Evonik Industries AG, Essen, Deutschland³ Institut für Arbeitsmedizin, Sozialmedizin und Umweltmedizin der Goethe-Universität Frankfurt/Main, Frankfurt/Main, Deutschland⁴ Europäische Forschungsvereinigung für Umwelt und Gesundheit im Transportsektor (EUGT e. V.), Berlin, Deutschland

Dieselmotoremissionen und Lungenkrebsrisiken

Bewertung der Metaanalyse von Vermeulen et al. 2014

Vermeulen et al. [44] publizierten eine Meta-Regressionsanalyse zu 3 wesentlichen epidemiologischen US-amerikanischen Studien [15, 36, 39], die den Zusammenhang zwischen Dieselmotoremissionen (DME, engl. DEE) – gemessen in der kumulierten Exposition gegenüber elementarem Kohlenstoff (EC) in $\mu\text{g}/\text{m}^3$ -Jahre – und der Lungenkrebsmortalität darstellten. Die Autoren beschreiben, dass einige von ihnen als Mitglieder der IARC-Working Group tätig waren [4], die 2012 ein Review und eine Neubewertung zur karzinogenen Wirkung von DME erstellte („hazard assessment“) und zur Einstufung von DME als Humankanzerogen führte (IARC Group 1). Vermeulen et al. [44] möchten mit dieser Folgearbeit die Überlegungen weiterführen und einen Beitrag zur quantitativen Risikoschätzung („risk assessment“) auf Basis der wesentlichen Studien leisten. Die Autoren ermittelten aus den Daten der 3 Studien als Hauptergebnis eine gemeinsame Expositions-Risiko-Kurve ohne Schwellenwert und folgerten: „We estimated a ln RR of 0,00098 (95 % CI 0,00055–0,0014) for lung cancer mortality with each 1- $\mu\text{g}/\text{m}^3$ -year increase in cumulative EC based on a linear meta-regression model. Corresponding ln RRs for the individual studies ranged from 0,00061 to 0,0012. Estimated numbers of excess lung cancer deaths

through 80 years of age for lifetime occupational exposures of 1, 10, and 25 $\mu\text{g}/\text{m}^3$ EC were 17, 200, and 689 per 10,000, respectively. For lifetime environmental exposure to 0.8 $\mu\text{g}/\text{m}^3$ EC, we estimated 21 excess lung cancer deaths per 10,000. (...) Our estimates suggest that stringent occupational and environmental standards for DEE should be set.“ Die Arbeit beschreibt somit erhöhte Krebsrisiken selbst bei sehr niedrigen DME-Belastungen. Der Arbeit von Vermeulen et al. [44] wird daher in den laufenden DME-Grenzwertdiskussionen eine wichtige Rolle zukommen. Das US-amerikanische National Institute for Occupational Health (NIOSH) hat nach seiner üblichen Vorgehensweise (Exzessfallberechnung) bereits eine Ableitung versucht [32]: Der Autor ermittelte auf Basis von Vermeulen et al. [44] eine Schätzung für den 8-h-Grenzwert für DME am Arbeitsplatz von 0,59 $\mu\text{g}/\text{m}^3$.

Da diese Grenzwertableitung deutlich niedriger liegt als beispielsweise die Empfehlung von 100 $\mu\text{g}/\text{m}^3$ DEEE („diesel engine exhaust emissions“), gemessen als „elemental carbon“, des ACSH der EU [42] und der SHEcan-Arbeitsgruppe [9], ist eine kritische Auseinandersetzung mit der Meta-Regressionsanalyse von Vermeulen et al. [44] angezeigt.

Material

Vermeulen et al. 2014: Inputdaten zur Hauptanalyse („Primary“)

Die Basisdaten der Metaanalysen von Vermeulen [mittlere Expositionsschätzung jeder Expositions-kategorie der eingeschlossenen Studien und jeweils zugehöriges relatives Risiko (RR) mit 95 % Konfidenzintervall] sind im Supplement von Vermeulen et al. [44] weitgehend enthalten. Diese Daten wurden extrahiert und in eine Stata-Datei übernommen. Lücken wurden gefüllt, soweit die fehlenden Angaben in den Originalpublikationen der Einzelstudien enthalten waren.

Als Beispiel zeigt **Tab. 1** die Grunddaten, die in die Hauptanalyse („Primary“) von Vermeulen et al. [44] eingegangen sind. Die 3 eingeschlossenen Studien [15, 36, 39] geben zu unterschiedlich gerasterter, aber einheitlich in $\mu\text{g}/\text{m}^3$ -Jahre gemessener kumulierter Exposition gegenüber DME als elementarem Kohlenstoff (Dosis) relative Risiken und die zugehörigen 95 %-Konfidenzintervall-Grenzen an.

Steenland et al. [39] ist eine genetete Fall-Kontroll-Studie zu Arbeitern in der US Trucking Industry (994 Lungenkrebstodesfälle und 1085 Kontrollen). Die Dosiswerte wurden um 5 Jahre „gelagt“, um

Tab. 1 Inputdaten zur Hauptanalyse („Primary“) in Vermeulen et al. [44]

Studie	Lag/a	Expositions-kategorie	Mittlere Dosis	Untere Dosis	Obere Dosis	RR	95 % CI		Anzahl Personen	Anzahl Fälle
							Unten	Oben		
Steenland et al. [39]	5	Referenz	0,0	0,0	0,0	1,00	1,00	1,00	–	–
	5	Kat 1	84,5	0,0	< 169,0	1,08	0,72	1,63	–	–
	5	Kat 2	231,0	169,0	257,0	1,10	0,74	1,65	–	–
	5	Kat 3	294,0	257,0	331,0	1,36	0,90	2,04	–	–
	5	Kat 4	551,7	≥ 331,0		1,64	1,09	2,49	–	–
Garshick et al. [15] ^a	5	Referenz	15,5	0,0	< 30,9	1,00	1,00	1,00	105.513	122
	5	Kat 1	51,3	30,9	71,7	1,31	1,01	1,71	104.909	191
	5	Kat 2	111,0	71,7	150,3	1,38	1,02	1,87	102.496	202
	5	Kat 3	250,5	≥ 150,3		1,48	1,05	2,10	87.397	226
Silverman et al. [36]	15	Referenz	1,5	0,0	< 3,0	1,00	1,00	1,00	207	49
	15	Kat 1	37,5	3,0	72,0	0,74	0,40	1,38	278	50
	15	Kat 2	204,0	72,0	536,0	1,54	0,74	3,20	206	49
	15	Kat 3	1036,0	≥ 536,0		2,83	1,28	6,26	173	50

Dosis bezeichnet die kumulierte Exposition gegenüber DME in $\mu\text{g}/\text{m}^3\text{-Jahre}$. Für jede Kategorie (Referenz, Kat 1, Kat 2, Kat 3 bzw. Kat 4) werden „Mittelwerte“ sowie die unteren als auch oberen Grenzen zur Dosis angegeben. Schätzer des relativen Risikos (RR) mit 95 % Konfidenzintervall (95 % CI) pro Kategorie. Umfang der Studien und Zahl der Lungenkrebstodesfälle.

^aGarshick et al. [15] enthält nicht die Jobgruppe „Mechaniker“ und die Risikoschätzer sind nach der Expositionsdauer adjustiert. Die Zahl der Personenjahre ist an Stelle der Personenzahl angegeben, da es sich um eine Kohortenstudie handelt. Steenland et al. [39] geben weder die Anzahl der Personen noch die Anzahl der Fälle pro Expositions-kategorie an.

Latenzzeitphänomene in der Berechnung der Odds-Ratios (OR) zu berücksichtigen. Mit „lagging“ wird eine Auswertetechnik bezeichnet, die Expositionswerte der letzten Jahre (hier: der letzten 5 Jahre) nicht in die Analyse einfließen lässt [33]. Diese Untersuchung ist die älteste der in die Analyse übernommenen Studien, da sie auf den Daten der Fall-Kontroll-Studie von Steenland et al. 1990 [38] beruht: „All cases and controls had died in 1982–1983“. Als Grundlage der Expositionsschätzung für den Zeitraum 1949 bis 1983 wurden die Messwerte zu „elemental carbon“ von Zarebst et al. 1991 [45] verwendet, die allerdings erst 1990 gewonnen wurden. Der Zeitpunkt der Messdatenerhebung liegt damit etwa 8 Jahre nach dem Tod der in die Studie eingeschlossenen Personen und entsprechend noch später nach Ende ihrer Expositionsphase. Steenland et al. [39] haben versucht, diese Daten rückwärts zu extrapolieren, „dependent on very broad assumptions“. Die Autoren bewerten diese wichtige Limitation ihrer Arbeit wie folgt: „Our results should be regarded with appropriate caution because our exposure estimates are based on broad assumptions rather than actual measurements“. Im Abstract schreiben sie: „Our results depend on estimates about unknown past

exposures, and should be viewed as exploratory.“

Garshick et al. [15] ist eine hiervon unabhängige Kohortenstudie zur US Trucking Industrie (31.135 männliche Arbeiter, 779 Lungenkrebstodesfälle). Vermeulen et al. [44] verwendeten diese Kohortendaten nach Ausschluss von Mechanikern und auch hier wurden die Dosiswerte um 5 Jahre gelagt. Es wurden jene Risikoschätzer (Hazard-Ratios, HR) der Garshick-Studie übernommen, die zusätzlich nach der Expositionsdauer adjustiert sind.

Silverman et al. [36] ist die Fall-Kontroll-Studie zur US-amerikanischen DEMS (Diesel Exhaust in Miners Study). Die zugrunde liegende Kohorte [3] umfasste 12.315 Bergarbeiter aus 8 Betrieben (kein Erz- oder Steinkohlenbergbau). Die Fall-Kontroll-Studie hat hiervon 198 Lungenkrebstodesfälle und 562 Kontrollen eingeschlossen. Die Autoren haben zur Berechnung der Odds-Ratios abweichend von den anderen beiden Studien einen lag von 15 Jahren verwendet.

Zur genaueren Beschreibung der eingeschlossenen Studien sei auf Vermeulen et al. [44] und auf die Originalpublikationen verwiesen. In **Tab. 1** sind die OR und HR einheitlich als relative Risiken aufgelistet.

Korrigierte Schätzer

In den Analysen von Vermeulen et al. [44] wurden stets die Risikoschätzer (HR) der Garshick-Studie übernommen, die zusätzlich nach der Expositionsdauer adjustiert worden waren. Ein wichtiger Aspekt hierzu ergibt sich aus dem Leserbrief von Morfeld [24] (inklusive Antwort der Autoren) zur Fehladjustierung der von Vermeulen et al. [44] verwendeten Koeffizienten. Morfeld kritisiert, dass die kumulierte Exposition zusätzlich nach der in ihr bereits enthaltenen Expositionsdauer adjustiert wird und damit der Risikokoeffizient nicht die Wirkung der kumulierten Exposition schätzt, sondern die einer Konzentration (obgleich dieser Zugang zu einer Schätzung der Konzentrationswirkung nicht optimal ist). Die Autoren antworteten auf die Kritik wie folgt: „Morfeld suggests that adjusting cumulative exposure by duration of employment time reduces cumulative exposure to an estimate of long-term average concentration. We agree that if exposure in our workers was relatively constant, cumulative exposure would be the simple product of duration and average exposure. However, exposure varies considerably over time and between and within jobs.“ Dieser Hinweis

P. Morfeld · M. Spallek

Dieselmotoremissionen und Lungenkrebsrisiken. Bewertung der Metaanalyse von Vermeulen et al. 2014**Zusammenfassung**

Einleitung. Vermeulen et al. 2014 publizierten eine Meta-Regressionsanalyse zu 3 wesentlichen epidemiologischen US-amerikanischen Studien zum Zusammenhang zwischen Lungenkrebsmortalität und beruflicher Dieselmotoremission (DME), die in der kumulierten Exposition gegenüber elementarem Kohlenstoff in $\mu\text{g}/\text{m}^3$ -Jahre gemessen wird. Die Gruppe um Vermeulen ermittelte eine statistisch signifikante Dosis-Wirkungs-Beziehung und erhöhte Krebsrisiken selbst bei sehr niedrigen Belastungen.

Methoden. Wir führten eine erweiterte Reanalyse durch.

Ergebnisse. Wir konnten die im Jahr 2014 berichteten Einzel- und Hauptbefunde zwar weitgehend reproduzieren, allerdings gab

es eine ausgeprägte Heterogenität im Ausgangsniveau. Eine Korrektur der Koeffizienten in Garshick et al. 2012 führte zu einer Reduktion dieser Heterogenität, bewirkte aber für diese Studie einen um eine Größenordnung kleineren Dosiskoeffizienten, weit entfernt von einer Signifikanz ($p=0,6$). Wird die Metaanalyse um eine Suche nach Schwellen erweitert, so ergibt sich mit korrigierten Garshick-Koeffizienten ein Schwellenwert für die kumulierte DME-Exposition bei $150 \mu\text{g}/\text{m}^3$ -Jahre (nicht signifikant).

Die von Vermeulen et al. 2014 durchgeführte Datenauswahl führte zu den höchsten Risikoschätzern. Der niedrigste Schätzer ergab sich nach Ausschluss der explorativen Studie von Steenland et al. 1998 in einer

Analyse der 3 Studien Garshick et al. 2012 (modifiziert), Silverman et al. 2012 (modifiziert nach Crump et al. 2015) und Möhner et al. 2013. Der Meta-Koeffizient sank auf etwa 10–20 % des Wertes, der von Vermeulen et al. 2014 als Hauptergebnis publiziert wurde.

Schlussfolgerung. Die Ergebnisse aus der Meta-Regressionsstudie von Vermeulen et al. 2014 sollten nicht vorbehaltlos in einer Risikoabschätzung verwendet werden. Dies gilt insbesondere für den Niedrigexpositionsbereich gegenüber DME.

Schlüsselwörter

Diesel · DME · Lungenkrebs · Epidemiologie · Metaanalyse

Diesel motor emission and lung cancer risks. Evaluation of the meta-analysis by Vermeulen et al. 2014**Abstract**

Introduction. Vermeulen et al. 2014 published a meta-regression analysis of three relevant epidemiological US studies. The investigations estimated the association between occupational diesel exhaust emission (DEE) and lung cancer mortality. The DEE dose was measured as a cumulative exposure to elemental carbon in $\mu\text{g}/\text{m}^3$ years. Vermeulen et al. 2014 found a statistically significant dose-response association and elevated lung cancer risks even at very low exposures.

Methods. We conducted an extended reanalysis.

Results. The present reanalysis could broadly reproduce the individual and main meta-analytical results; however, the analysis dem-

onstrated a heterogeneity of the offsets between the studies. This heterogeneity was reduced after the coefficients of Garshick et al. 2012 were corrected while the meta-coefficient for dose dropped by an order of magnitude for this study and was far from being significant ($p=0,6$). A (non-significant) threshold estimate for the cumulative DEE exposure was found at $150 \mu\text{g}/\text{m}^3$ years when extending the meta-analyses of the three studies by hockey stick regression modeling (using the corrected coefficients of Garshick et al. 2012).

The data used by Vermeulen et al. 2014 led to the highest risk estimates. The lowest estimate was found after exclusion of the ex-

plorative study of Steenland et al. 1998 in an analysis of Garshick et al. 2012 (modified), Silverman et al. 2012 (modified according to Crump et al. 2015) and Möhner et al. 2013. The meta coefficient was found to be approximately 10–20 % of the main effect estimate reported by Vermeulen et al. 2014.

Conclusion. The findings of Vermeulen et al. 2014 should not be used without reservations in any risk assessments. This is particularly true for the low end of the exposure scale.

Keywords

Diesel · DEE · Lung cancer · Epidemiology · Meta-analysis

rechtfertigt die Vorgehensweise aber dennoch nicht, denn in den durchgeführten Cox-Analysen werden die Daten zeitabhängig ausgewertet. Damit gilt für jeden Zeitpunkt und in jeder Person: kumulierte Exposition = Expositionszeit × Durchschnittskonzentration. Durch die Adjustierung für die Expositionszeit wird die Auswertung dahingehend verändert, dass die zeitabhängige Durchschnittskonzentration und nicht die kumulierte Exposition analysiert wird. Die Autoren begründen die von ihnen dennoch durchgeführte Adjustierung nach der Expositionszeit mit einer angeblich erreichten Adjustie-

rung für den Healthy Worker Survivor Effect. Hierfür sind jedoch andere Methoden notwendig [33].

In keiner der anderen beiden Studien [36, 39] wurde eine solche Adjustierung der kumulierten Exposition für die Expositionszeit durchgeführt. Ein weiterer Aspekt zur Garshick-Studie, der ebenfalls den Modell Aufbau betrifft, wird im Diskussionskapitel besprochen.

Koeffizienten zur kumulierten Exposition können Garshick et al. [15] entnommen werden, auch ohne dass gleichzeitig nach der Expositionszeit adjustiert wurde. Um den Koeffizienten ihren übli-

chen Sinn geben zu können und um die Vergleichbarkeit mit den anderen beiden Studien zu verbessern, wurden diese Koeffizienten zusätzlich in diese Analyse übernommen. Die Daten sind in **Tab. 2** dargestellt und werden im Folgenden als Garshick et al. [15] (modifiziert) bezeichnet.

Die Risikoschätzer sind nicht nur niedriger als in **Tab. 1**, sondern zeigen auch mit zunehmender Expositionszeit keinen positiven Trend an.

Vermeulen et al. [44] führten auch Sensitivitätsanalysen mit variierten Daten zu Garshick et al. [15] durch. Allerdings be-

Tab. 2 Ergebnisse von Garshick et al. [15] ohne Adjustierung nach Expositionsdauer, als Garshick et al. [15] (modifiziert) bezeichnet

Expositions-Kategorie	Mittlere Dosis	Untere Dosis	Obere Dosis	RR	95 % CI		Personenjahre	Anzahl Fälle
					Unten	Oben		
Referenz	15,5	0,0	< 30,9	1,00	1,00	1,00	105.513	122
Kat 1	51,3	30,9	71,7	1,18	0,92	1,52	104.909	191
Kat 2	111,0	71,7	150,3	1,17	0,88	1,55	102.496	202
Kat 3	250,5	≥ 150,3		1,19	0,86	1,63	87.397	226

Dosis bezeichnet die kumulierte Exposition gegenüber DME in $\mu\text{g}/\text{m}^3\text{-Jahre}$. Für jede Kategorie (Referenz, Kat 1, Kat 2, Kat 3) werden Mittelwerte sowie die unteren als auch oberen Grenzen zur Dosis angegeben. Odds-Ratio (OR) mit 95 % Konfidenzintervall (95 % CI) sowie Anzahl der Personenjahre unter Beobachtung und Anzahl der Lungenkrebstodesfälle pro Dosiskategorie.

Tab. 3 Reanalyse zu Silverman et al. [36] mit Adjustierung für die Radonexposition. (Tab. III in Crump et al. [13])

Expositionskategorie	Mittlere Dosis	Untere Dosis	Obere Dosis	OR	95 % CI unten oben		Anzahl Personen	Anzahl Fälle
Crump et al. [13], REC4								
Referenz	0,71	0	< 4,9	1	1	1	217	49
Kat 1	26,65	4,9	< 70,4	0,80	0,41	1,55	266	50
Kat 2	243,43	70,4	< 498,4	1,67	0,73	3,81	192	49
Kat 3	1522,10	–	≥ 498,4	1,50	0,54	4,17	189	50
Crump et al. [13], REC6								
Referenz	0,61	0	< 2,8	1	1	1	230	49
Kat 1	20,47	2,8	< 50,6	1,07	0,57	2,00	247	50
Kat 2	158,27	50,6	< 388,0	1,35	0,62	2,94	206	49
Kat 3	1156,89	–	≥ 388,0	1,43	0,52	3,94	181	50

Die Ergebnisse beruhen auf den Expositionsmetriken REC4 und REC6 mit einem lag von 15 Jahren. Dosis bezeichnet die kumulierte Exposition gegenüber DME in µg/m³-Jahre. Für jede Kategorie (Referenz, Kat 1, Kat 2, Kat 3) werden Mittelwerte sowie die unteren als auch oberen Grenzen zur Dosis angegeben. Odds-Ratio (OR) mit 95 % Konfidenzintervall (95 % CI) sowie Anzahl der Personen unter Beobachtung und Anzahl der Lungenkrebstodesfälle pro Dosiskategorie.

treffen diese Variationen allein den „lag“ der Expositionsvariablen und den Ein- bzw. Ausschluss der Daten der Fallgruppe „Mechaniker“. Das Problem der nach Expositionsdauer fehladjustierten Koeffizienten der kumulierten Exposition wird von den Autoren hingegen nicht berücksichtigt.

Reanalyse der DEMS-Fall-Kontroll-Studie: Crump et al. 2015

Crump et al. [13] reanalysierten die DEMS-Fall-Kontroll-Studie von Silverman et al. [36] und konnten deren Ergebnisse weitgehend reproduzieren: „We were able to replicate the findings reported by Silverman et al. (18) when we used the same analytical methods. This gave us confidence that we were using the same basic data set as Silverman et al.“ Crump

und Mitarbeiter untersuchten den Einfluss von Kovariablen, die Silverman et al. [36] nicht in ihre finalen Modelle übernommen hatten. Als wesentlichster Confounder stellte sich die Radon-Exposition unter Tage heraus, ein Ergebnis, dass sich nicht mit den Aussagen von Silverman et al. [36] deckt. Silverman et al. [36] schrieben zur kumulierten Exposition gegenüber Radon „estimated cumulative exposure to radon (...) were evaluated but not included in the final models because they had little or no impact on odds ratios (i.e., inclusion of these factors in the final models changed point estimates for diesel exposure by $\leq 10\%$).“ Crump et al. [13] bemerkten hierzu: „However, when we reproduced the Silverman et al. analysis, we could not verify this statement.“ Im Folgenden werden daher in der vorliegenden Sensitivitätsanalyse die Schätzungen von

Crump et al. [13] ausschließlich nach einer zusätzlichen Adjustierung für die kumulierte Radonexposition verwendet.

Crump et al. [13] entwickelten zudem 6 neue DME-Expositionsmaße, alternativ zu den Schätzungen, die in Attfield et al. [3] und Silverman et al. [36] eingesetzt wurden: „We proceeded to apply six alternative REC [respirable elemental carbon] metrics, five of which depended, as did the DEMS metrics, on extrapolations involving assumed relationships between CO [carbon monoxide] and REC. A sixth REC metric, REC6, was used that did not involve any assumptions concerning the relationship between CO and REC, and was based on Adj_HP [adjusted horse power] and ventilation rates for each of the mines. Of the several REC metrics, we view REC6 as having some superior qualities because it avoids using the highly uncertain assumptions concerning the relationship between CO and REC.“ Wir verwenden daher REC6 als Hauptalternative zu den Silverman-Expositionsdaten.

In ihren zusätzlichen Auswertungen zur Expositionsschätzung ermittelten Crump und van Landingham [11] als besten Schätzwert $\beta=0,3$ im Umrechnungsmodell $\text{REC} \sim \text{CO}^\beta$, im Unterschied zu dem Wert von $\beta=0,58$ nach Stewart et al. [40] oder zu dem Wert $\beta=1,0$, der von Silverman et al. [36] in ihren Auswertungen angenommen wurde. Die alternative Expositionsmetrik REC4 [13] enthält diesen besten Schätzwert $\beta=0,3$. REC4 ist zudem direkt auf die Arbeiten zur alternativen Expositionsschätzung von Crump und van Landingham [11] gegründet und berücksichtigt einige weitere Aspekte (vgl. die detaillierte Beschreibung zu REC1 bis REC 4 in [13]). Wir verwenden daher zusätzlich die Expositionsmetrik REC4 als zweite Variation nach Crump et al. [13]. Einen Überblick zu den Ergebnissen der Reanalyse mit REC4 und REC6 nach Crump und Mitarbeitern zur Silverman-Studie gibt Tab. 3.

Neu kategorisierte Daten: Möhner et al. 2013 (adaptiert)

Die deutsche Kalibergbaustudie [21] wurden von Vermeulen et al. [44] von der eigentlichen Analyse ausgeschlossen, da die Referenzkategorie zu hoch definiert

Tab. 4 Zusätzliche Ergebnisse zur Studie von Möhner et al. [21] mit geänderter Kategorisierung, bezeichnet als Möhner et al. [21] (adaptiert)

Expositions-kategorie	Mittlere Dosis	Untere Dosis	Obere Dosis	OR	95 % CI		Anzahl Personen	Anzahl Fälle
					Unten	Oben		
Referenz	263	0	≤ 380	1	1	1	29	5
Kat 1	728	380	1024	0,99	0,323	3,031	126	19
Kat 2	1386	1024	1840	1,69	0,53	5,403	126	25
Kat 3	2538	> 1840		1,2	0,363	3,971	127	19

Dosis bezeichnet die kumulierte Exposition gegenüber DME in $\mu\text{g}/\text{m}^3\text{-Jahre}$. Für jede Kategorie (Referenz, Kat 1, Kat 2, Kat 3) werden Mittelwerte sowie die unteren als auch oberen Grenzen zur Dosis angegeben. Odds-Ratio (OR) mit 95 % Konfidenzintervall (95 % CI) sowie Anzahl der Personen unter Beobachtung und Anzahl der Lungenkrebstodesfälle pro Dosiskategorie.

worden sei. Nur in 2 Sensitivitätsanalysen wurde die Arbeit von Möhner et al. [21] berücksichtigt, allerdings indem Vermeulen et al. [44] die RR-Schätzer entweder original beließen oder eine Ad-hoc-Korrektur der Risikoschätzer vornahmen. Beide Zugänge sind nicht optimal, weshalb keine dieser Auswertungen von Vermeulen et al. [44] im Rahmen der Hauptanalyse dargestellt, sondern zu den Ergebnissen lediglich im Online-Appendix berichtet wird.

Um die deutsche Kalibergbaustudie in diesem Projekt möglichst informativ berücksichtigen zu können, mussten die Daten in anderer Rasterung übernommen werden als publiziert. Mit dem Erstautoren wurde Kontakt aufgenommen, um die hierzu notwendigen Angaben zu bekommen (die Publikation enthält nicht alle interessierenden Rasterungen der Exposition, und es fehlen weitere Detailangaben). Auf Anfrage übergab Dr. Möhner per E-Mail am 03.07.2014 zusätzliche Ergebnisse zur Studie von Möhner et al. [21]. Diese stellten Fallzahlen und Odds-Ratio-Schätzungen mit 95 % Konfidenzintervallen zu einer geänderten Kategorisierung dar, die im Folgenden als „Möhner et al. [21] (adaptiert)“ bezeichnet werden.

Die Aufteilung wurde so adaptiert, dass die Referenzkategorie nur noch mit 5 Fällen besetzt ist. Somit liegt die Expositionshöhe der Referenzkategorie deutlich niedriger als in der Originalanalyse (Kritikpunkt in [44]). Die anderen 3 Kategorien wurden so gewählt, dass sich die Fälle annähernd gleich verteilen. Die Ergebnisse dieser neuen Kategorisierung von Möhner et al. [21] zeigt **Tab. 4**.

Methoden

Reproduktion der Ergebnisse und Variation der Modellierung

Die komplexe Metaanalyse der Autoren zu den kategorisierten Einzelstudien-ergebnissen wird nachvollzogen. Eine Herausforderung stellt die mit den Studien variierende Kategorisierung der Exposition dar und potenzielle Unterschiede in Design und Umfang der Studien. Zudem sind die Ergebnisse zu den Expositionskategorien innerhalb der Studien geschachtelt, so dass eine Zwei-Ebenen-Struktur vorliegt (1. Ebene: Expositionsgruppen, 2. Ebene: Studien). Um diese komplexe Datenlage angemessen analysieren zu können, werden zu den Ergebnissen der Einzelstudien sog. Meta-Regressionsanalysen sowohl mit festen [2] als auch mit zufälligen [6, 7, 14] Effekten durchgeführt, um eine gemeinsame Expositions-Wirkungs-Kurve zu ermitteln [5].

Über die Varianzen der RR-Schätzer pro Expositionskategorie müssen Gewichte in die Analysen einfließen. Diese Wichtung ist notwendig, da nur aggregierte Daten (also keine Einzeldaten zu Personen) in den Auswertungen zur Verfügung stehen und den Datenpunkten zu den Expositionskategorien aufgrund der variierenden Fallzahl innerhalb und zwischen den Studien eine unterschiedliche Relevanz in einer gemeinsamen Auswertung zukommt. Hierzu werden verschiedene Ansätze erprobt:

- lineare Regression mit festen Effekten zu log RR mit Gewichten proportional zum Inversen der jeweiligen Varianz
 - ohne Adjustierung nach Studien,

- mit Adjustierung nach Studien. Ein globaler F-Test zur Heterogenität der Risikoniveaus (Achsenabschnitte, Offsets) zwischen den Studien wird berechnet.
- gemischte lineare Regression zu log RR mit einem zufälligen Achsenabschnitt, der Unterschiede zwischen den Studien aufnimmt
 - mit Gewichten auf der ersten Ebene (Ebene der Expositions-kategorien) proportional zum Inversen der jeweiligen Varianz und mit den Summen dieser Gewichte als Studiengewichte auf der zweiten Ebene (Studienebene),
 - in einem zweiten Auswertungsansatz werden die Gewichte der ersten Ebene effektiv auf die Gewichte der zweiten Ebene skaliert.
- gemischte lineare Regression zu log RR mit einem zufälligen Achsenabschnitt und einem zufälligen Dosiskoeffizienten („slope“), die Unterschiede zwischen den Studien aufnehmen
 - mit Gewichten auf der ersten Ebene (Ebene der Expositions-kategorien) proportional zum Inversen der jeweiligen Varianz und mit den Summen dieser Gewichte als Studien-Gewichte auf der zweiten Ebene (Studienebene),
 - in einem zweiten Auswertungsansatz werden die Gewichte der ersten Ebene effektiv auf die Gewichte der zweiten Ebene skaliert.

Die Referenzpunkte werden in den Analysen nicht berücksichtigt, da ihr Gewicht keine reelle Zahl darstellt, sondern unendlich ist. Dies entspricht dem Vorgehen von Vermeulen et al. [44]: Fig. 1 der Studie enthält nicht 13, sondern lediglich 10 Beobachtungspunkte.

Modelle mit festen Effekten werden üblicherweise über die Student-t-Verteilung statistisch bewertet [28]. Da nur wenige Datenpunkte eingehen, die keine Individuen sondern allein die Gruppen wiedergeben (aggregierte Daten), führt dies tendenziell zu einer Überschätzung der p-Werte und Konfidenzintervallbreiten. Präzisionsgewichtungen bringen jedoch lediglich die relativen Gewichtsunterschiede zwischen den Datenpunk-

ten in die Analysen ein und lösen dieses Problem nicht. Alternativ werden daher die Modelle mit festen Effekten nach der Standardnormalverteilung bewertet [28]. Dies führt tendenziell zu p-Werten und Konfidenzintervallen, die zu eng sind, insbesondere da die Korrelationen zwischen den Gruppen innerhalb der Studien nicht berücksichtigt werden. In allen gemischten Regressionen erfolgt die Bewertung grundsätzlich über die Standardnormalverteilung.

Die zu den verschiedenen Expositions-kategorien berichteten Ergebnisse einer Studie sind allerdings nicht unabhängig voneinander, sondern korrelieren miteinander, da sie sich innerhalb einer Studie auf eine gemeinsame Referenzkategorie beziehen. Deshalb werden ergänzend zu den oben beschriebenen Auswertungen auch Meta-Regressionsverfahren mit Berücksichtigung der Korrelationen der Studienergebnisse innerhalb der Studien versucht [16, 30, 31]. Jedoch ist zum Einsatz dieses Greenland/Longnecker-Verfahrens zusätzlicher Input notwendig: Es müssen Personenjahre und Fallzahlen bzw. Personen- und Fallzahlen je Expositions-kategorie der Studien verfügbar sein. Diese Daten sind nicht im Supplement von Vermeulen et al. [44] enthalten und können auch nicht vollständig aus den Originalpublikationen rekonstruiert werden.

Im Methodenteil erwähnen Vermeulen et al. [44] fast alle diese Verfahren, aber es bleibt z. B. unklar, ob die Greenland/Longnecker-Methode stets zum Einsatz kam und welche Gewichtsstruktur in den gemischten Regressionen verwendet wurde. Mit unterschiedlichen Auswertungsansätzen soll in diesem Forschungsprojekt geprüft werden, in welchem Umfang und mit welcher Methode die publizierten Ergebnisse von Vermeulen et al. [44] reproduziert werden können.

Vermeulen et al. [44] setzten zudem Spline-Modelle ein, die eine flexiblere Kurvenform erlauben als log-lineare Modelle. Allerdings ergaben diese Spline-Regressionen keinen abweichenden Verlauf der geschätzten Expositions-Wirkungs-Beziehung: „The linear model (Fig. 1) and the spline meta-regression model (data not shown) fit the data well, with virtually equivalent curves.“ Wir setzen deshalb auch keine Spline-Funktionen oder ähn-

liche Verfahren ein, sondern folgen dem Hauptansatz von Vermeulen und Mitarbeitern, auch um die Zahl der zu schätzenden Parameter niedrig zu halten.

Als zentraler Modellierungsansatz soll aufgrund der höheren Stabilität und einfacheren Interpretation eine präzisionsgewichtete Regressionsanalyse mit festen Effekten bei gleichzeitiger Adjustierung für die Einzelstudien verfolgt werden. Dieser Auswertungsansatz hat zudem weitere Vorteile gegenüber einer Regression mit zufälligen Effekten (vgl. die Argumente in [2], S. 2, 3 und [6], S. 700). Es wird untersucht, ob die anderen Ansätze relevant verschiedene Ergebnisse hervorbringen. Falls das nicht der Fall ist, wird diese Methodik als führende Analysestrategie verwendet.

Alle Analysen werden mit Stata 12 durchgeführt [37].

Einfluss der gewählten Inputdaten

Ein wichtiger Aspekt zur Datenauswahl ergibt sich aus dem Leserbrief von Morfeld [24] zur Arbeit von Garshick et al. [15] mit dem Hinweis, dass die von Vermeulen et al. [44] verwendeten Koeffizienten fehladjustiert sind. In einer weiteren Analyse werden daher die oben beschriebenen Metaauswertungen mit den korrigierten Koeffizienten nach Garshick et al. [15] (modifiziert) wiederholt (■ Tab. 2).

Ergebnisse zur DEMS-Reanalyse wurden in einem HEI (Health Effects Institute) Webinar präsentiert [12, 22] und im Journal *Risk Analysis* publiziert [13, 23]. Crump et al. [13] enthält überarbeitete OR-Schätzungen zu Silverman et al. [36]; wichtige alternative Schätzungen aus dieser Arbeit sind in ■ Tab. 3 dargestellt. Die „REC6“-Befunde aus dieser Arbeit werden nach Adjustierung für die Radonexposition in der Metaanalyse an Stelle der Daten von Silverman et al. [36] (■ Tab. 1) berücksichtigt. Des Weiteren wird untersucht, inwieweit die Ergebnisse variieren, wenn die „REC4“-Befunde verwendet werden.

Die deutsche Kalibergbaustudie [21] soll im Sinne einer Sensitivitätsanalyse zusätzlich eingeschlossen werden. Wir verwenden die Risikoschätzer aus „Möhner et al. [21] (adaptiert)“ (■ Tab. 4).

Da die Arbeit von Steenland et al. [39] erhebliche Limitationen aufweist, werden

auch Metaanalysen durchgeführt, ohne diese Studie einzuschließen.

Nichtlinearität: Schwellenwertsuche

Die Daten werden nach Nichtlinearitäten in der Expositions-Risiko-Beziehung untersucht, wobei insbesondere eine systematische Suche nach Schwellenwerten für die kumulierte Exposition erfolgt [18, 19, 43]. Hierzu wird der in Morfeld et al. [27] ausführlich dargestellte Algorithmus zur Ermittlung eines No-Adverse-Effect-Ausgangsplateaus in der kumulierten Exposition mit den Meta-Regressionsverfahren kombiniert. Diese Analyse beschränkt sich auf den zentralen Modellierungsansatz (präzisionsgewichtete Regressionsanalyse mit festen Effekten bei gleichzeitiger Adjustierung für die Einzelstudien).

Als Schrittweite der Schwellenwert-exploration wird 10 µg/m³-Jahre gewählt (Bereich: 0 µg/m³-Jahre bis 500 µg/m³-Jahre). Somit werden pro Schwellenwert-suche 51 Modelle berechnet und verglichen. Die Schwellenwertanalyse wird mit Sonderprogrammen in Stata 12 durchgeführt [37].

Ergebnisse

Reproduktion der Ergebnisse in Vermeulen et al. 2014 und erweiterte Analysen

Vermeulen et al. [44] berichten in ■ Tab. 1 (Hauptanalyse, „Primary“) zu den 3 Studien von Steenland et al. [39], Garshick et al. [15] und Silverman et al. [36]. Die Autoren berichten Einzelbefunde (ermittelte Risikokoeffizienten) und das Ergebnis der Metaanalyse der Risikokoeffizienten. Im Folgenden werden diese Angaben mit den Ergebnissen der Reanalyse verglichen.

Einzelanalysen zu den Studien

Alle Einzelanalysen werden mit linearen Regression mit festen Effekten zu log RR nachgerechnet (aus methodischen Gründen fällt diese Analyse zu einer einzigen Studie mit der gemischten Regression zusammen). In den ■ Tab. 5, 6, 7 und 8 sind die Ergebnisse zu Steenland et al. [39],

Tab. 5 Steenland et al. [39]: Lineare Regression mit festen Effekten zu log RR mit Gewichten proportional zum Inversen der jeweiligen Varianz

Log RR	Koef.	SE	t	P	95 % CI	
Dosis	0,00096	0,00021	4,55	0,045 ($< 0,001$)	0,00005 (0,00055)	0,00187 (0,00137)
Konstante	−0,031	0,070	−0,45	0,70	−0,33	0,27

Koeffizient (Koef.) der kumulierten Exposition (Dosis) gegenüber DME in $\mu\text{g}/\text{m}^3$ -Jahre und Absolutterm (Konstante) mit Standardabweichung (SE); t-Teststatistik zum Koeffizienten mit zugehörigem p-Wert und 95 % Konfidenzintervall zum Koeffizienten (95 % CI). In Klammern die Ergebnisse bei Anwendung der Standardnormalverteilung.

Tab. 6 Garshick et al. [15]: Lineare Regression mit festen Effekten zu log RR mit Gewichten proportional zum Inversen der jeweiligen Varianz

Log RR	Koef.	SE	t	p-Wert	95 % CI	
Dosis	0,00061	0,000091	6,6	0,095 ($< 0,001$)	−0,00055 (0,00042)	0,00177 (0,00078)
Konstante	0,244	0,013	18,6	0,034	0,078	0,411
<i>Garshick (modifiziert)</i>						
Dosis	0,00005	0,00007	0,75	0,59 (0,45)	−0,00077 (−0,00008)	0,00087 (0,00017)
Konstante	0,159	0,009	16,87	0,038	0,039	0,279

Koeffizient (Koef.) der kumulierten Exposition (Dosis) gegenüber DME in $\mu\text{g}/\text{m}^3$ -Jahre und Absolutterm (Konstante) mit Standardabweichung (SE); t-Teststatistik zum Koeffizienten mit zugehörigem p-Wert und 95 % Konfidenzintervall zum Koeffizienten (95 % CI). In Klammern die Ergebnisse bei Anwendung der Standardnormalverteilung. Ausgewertet werden Garshick et al. [15], wie von Vermeulen et al. [44] verwendet, und Garshick et al. [15] (modifiziert).

Tab. 7 Silverman et al. [36]: Lineare Regression mit festen Effekten zu log RR mit Gewichten proportional zum Inversen der jeweiligen Varianz

Log RR	Koef.	SE	t	p-Wert	95 % CI	
Dosis	0,00121	0,00055	2,20	0,27 (0,028)	−0,00579 (0,00013)	0,00821 (0,00229)
Konstante	−0,148	0,299	−0,49	0,71	−3,95	3,65
<i>Crump et al. [13], REC4</i>						
Dosis	0,00033	0,00051	0,64	0,63 (0,53)	−0,00615 (−0,00067)	0,00680 (0,00132)
Konstante	0,00541	0,358	0,02	0,99	−4,54	4,55
<i>Crump et al. [13], REC6</i>						
Dosis	0,00021	0,00021	1,01	0,50 (0,31)	−0,00247 (−0,00020)	0,00289 (0,00063)
Konstante	0,137	0,108	1,27	0,42	−1,23	1,51

Koeffizient (Koef.) der kumulierten Exposition (Dosis) gegenüber DME in $\mu\text{g}/\text{m}^3$ -Jahre und Absolutterm (Konstante) mit Standardabweichung (SE); t-Teststatistik zum Koeffizienten mit zugehörigem p-Wert und 95 % Konfidenzintervall zum Koeffizienten (95 % CI). In Klammern die Ergebnisse bei Anwendung der Standardnormalverteilung. Ausgewertet wird Silverman et al. [36], wie von Vermeulen et al. [44] verwendet, und entsprechend der Reanalyse von Crump et al. [13] mit Adjustierung für die Radonexposition (REC4, REC6; Tab. 3).

Tab. 8 Möhner et al. [21] (adaptiert): Lineare Regression mit festen Effekten zu log RR mit Gewichten proportional zum Inversen der jeweiligen Varianz

Log RR	Koef.	SE	t	p-Wert	95 % CI	
Dosis	0,00007	0,00029	0,24	0,85 (0,81)	−0,00364 (−0,00050)	0,00378 (0,00064)
Konstante	0,119	0,492	0,24	0,85	−6,13	6,37

Koeffizient (Koef.) der kumulierten Exposition (Dosis) gegenüber DME in $\mu\text{g}/\text{m}^3$ -Jahre und Absolutterm (Konstante) mit Standardabweichung (SE); t-Teststatistik zum Koeffizienten mit zugehörigem p-Wert und 95 % Konfidenzintervall zum Koeffizienten (95 % CI). In Klammern die Ergebnisse bei Anwendung der Standardnormalverteilung.

Garshick et al. [15], Silverman et al. [36] und Möhner et al. [21] dargestellt.

Folgende Koeffizientenschätzungen werden in Tab. 1 in Vermeulen et al. [44] zu Steenland et al. [39] berichtet: Dosis = 0,00096 (95 % CI 0,00033 bis 0,00159), Konstante = −0,032. Die Übereinstimmung in den Koeffizienten ist sehr gut. Das Konfidenzintervall ist stimmig, wenn beide Berechnungsmethoden (Student-t-Verteilung, Normalverteilung) berücksichtigt werden.

Folgende Koeffizientenschätzungen werden in Tab. 1 in Vermeulen et al. [44] zu Garshick et al. [15] berichtet: Dosis = 0,00061 (95 % CI −0,00088 bis 0,00210), Konstante: 0,24. Auch hier stimmt die Koeffizientenschätzung sehr gut und die Intervallschätzung weitgehend überein. Wird Garshick et al. [15] (modifiziert) ausgewertet, so ergeben sich deutlich andere Schätzungen: Der Dosiskoeffizient ist mehr als eine Zehnerpotenz kleiner und weit entfernt von einer Signifikanz ($p > 0,4$). Auffällig ist ein signifikantes Abweichen der Konstanten (Achsenabschnitt) vom Normalwert Null. Dieser Normalwert Null bedeutet ein Ausgangsniveau des relativen Risikos von 1, wie es bei einer DME-Exposition von 0 $\mu\text{g}/\text{m}^3$ -Jahre aber zu unterstellen ist.

Die folgenden Koeffizientenschätzungen werden in Tab. 1 in Vermeulen et al. [44] zu Silverman et al. [36] berichtet. Dosis = 0,00120 (95 % CI: 0,00053 bis 0,00187), Konstante: −0,18. Auch hier stimmen die Koeffizientenschätzungen gut überein (vermutlich liegt bei [44] ein Schreibfehler vor: −0,18 statt −0,148). Die Intervallschätzungen sind ebenfalls kompatibel.

Die Dosis-Koeffizienten der Crump-Variationen hingegen sind deutlich kleiner als die Originalwerte: Sie betragen lediglich 27 % (REC4) bzw. 18 % (REC6) des Silverman-Koeffizienten. Die t-Test-Konfidenzintervalle liegen in diesen Rechnungen weitgehend symmetrisch um die Null, so dass beide Modelle keinen Hinweis auf einen Trend ergeben. Auch bei Verwendung der Standardnormalverteilung ergibt sich kein Hinweis auf einen signifikanten Einfluss der DME-Exposition auf die Lungenkrebsmortalität.

In Tab. 8 ist das Ergebnis der Auswertung der Studie von Möhner et al. [21]

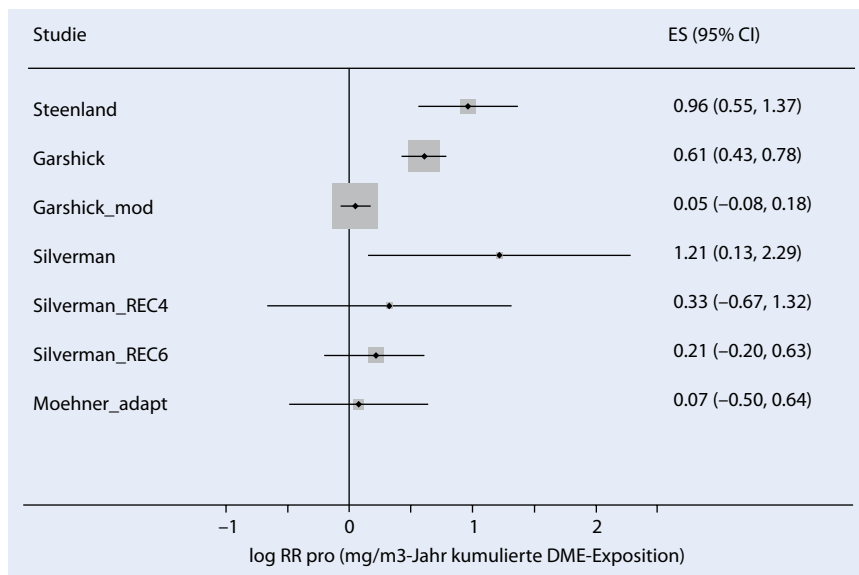


Abb. 1 ▲ Analyse der Einzelstudien. Lineare Regression mit *festen* Effekten zu log RR mit Gewichten proportional zum Inversen der jeweiligen Varianz (dargestellt durch Boxgröße). Koeffizienten („effect size“, ES) der kumulierten Exposition (Dosis) gegenüber DME in mg/m³-Jahre mit 95 % Konfidenzintervall zum Koeffizienten, berechnet mit Hilfe der Standardnormalverteilung. Ausgewertet werden Steenland et al. [39] und Silverman et al. [36], wie von Vermeulen et al. [44] verwendet, und entsprechend der Reanalyse von Crump et al. [13] mit Adjustierung für die Radonexposition (REC4, REC6; Tab. 3), Garshick et al. [15], wie von Vermeulen et al. [44] verwendet, und Garshick et al. [15] (modifiziert) sowie die Daten nach Möhner et al. [21] (adaptiert)

mit geänderter (adaptierter) Kategorisierung der Exposition dargestellt.

Die deutsche Studie ergibt keinen Hinweis auf einen Zusammenhang zwischen DME-Exposition und Lungenkrebsmortalität.

■ **Abb. 1** enthält einen Überblick zu den Ergebnissen der Regressionen mit festen Effekten bei Annahme einer Normalverteilung zur statistischen Bewertung für alle Einzelstudien. Modelle mit festen Effekten werden dargestellt, um die Koeffizienten nicht zu niedrig einzuschätzen. Die Normalverteilung wurde gewählt, um eine Überschätzung der p-Werte auszuschließen (d. h. die dargestellten Konfidenzintervalle sind keinesfalls zu weit). Abweichend von Vermeulen et al. [44] wird als Expositionseinheit mg/m³-Jahre = 1000 × µg/m³-Jahre gewählt, um die Darstellung übersichtlich zu halten. Jeder der 3 Datensätze, der in die Vermeulen-Analyse einging [15, 36, 39], ergibt einen signifikant erhöhten Risikoschätzer, wobei die Arbeit von Garshick et al. [15] den größten Einfluss hat. Werden die Ergebnisse von Garshick et al. [15] ohne Adjustierung nach Expositionsdauer, also Garshick et al. [15] (modifiziert) analysiert, so

sinkt der Risikoschätzer beträchtlich und ist nicht signifikant. Auch die Crump-Modifikationen (REC4, REC6) der Arbeit von Silverman et al. [36] ergeben deutlich niedrigere Risikoschätzer als die Originalarbeit und zeigen keine statistische Signifikanz. Die Studie von Möhner et al. [21] gibt keinen Hinweis auf einen Zusammenhang zwischen DME-Exposition und Lungenkrebsmortalität.

Gemeinsame Analysen der Studien

Laut ■ **Tab. 1** in Vermeulen et al. [44] ergeben sich für „all studies combined“ (d. h. [15, 36, 39]) als Koeffizienten die kumulierte Exposition (Dosis) gegenüber DME in µg/m³-Jahre und den Absolutterm (Konstante) folgende Werte: Dosis = 0,00098 (95 % CI 0,00055 bis 0,00141) und Konstante = 0,088. ■ **Tab. 9** berichtet zu der Reanalyse der 3 Studien mit Hilfe einer log-linearen Regression mit festen Effekten, ohne nach den Studien zu adjustieren.

In ■ **Tab. 9** ergeben sich relevante Abweichungen von den Ergebnissen in Vermeulen et al. [44]. Zwar ist der Expositions-Respons-Zusammenhang statistisch signifikant, aber der ohne Adjustierung

für die Studien geschätzte Dosiskoeffizient ist um etwa 20 % niedriger als von Vermeulen und Mitarbeitern publiziert (0,00076/0,00098 = 0,77). Werden die Versionen der Ergebnisse nach Crump et al. [13] in die Auswertung einbezogen, so ergibt die Metaanalyse keinen signifikanten Zusammenhang zwischen DME-Exposition und der Lungenkrebsmortalität.

In ■ **Tab. 10** werden die Koeffizientenschätzungen zusätzlich nach den Studien adjustiert.

■ **Tab. 10** ergibt für den Dosiskoeffizienten und das zugehörige Konfidenzintervall eine gute Übereinstimmung mit den in Vermeulen et al. [44] berichteten Ergebnissen (Dosis = 0,00098, 95 % CI 0,00055 bis 0,00141; Konstante = 0,088). Der nachgerechnete Koeffizient weicht minimal nach oben ab. Die beiden Konfidenzintervallberechnungen (Student-t-Verteilung, Normalverteilung) ergeben weitgehend übereinstimmende Ergebnisse, so dass die 10 Datenpunkte bereits ausreichen, um eine robuste Schätzung zu ermöglichen, die zudem mit den von Vermeulen und Mitarbeitern veröffentlichten Angaben übereinstimmt. Der zusätzlich ermittelte globale F-Test zur Heterogenität der Risikoniveaus zwischen den Studien ergibt $F(2,6) = 5,5$ ($p = 0,044$). Die Heterogenität zwischen den 3 Studien ist somit auf dem 5 %-Niveau signifikant. Mit Heterogenität ist ein systematischer Unterschied im Ausgangsrisiko (d. h. ohne Berücksichtigung der Exposition) der 3 Studien gemeint. Um die 3 Studien sinnvoll kombinieren zu können, sollten sie im Ausgangsniveau bis auf zufällige Abweichungen übereinstimmen. Jedoch sind die Unterschiede statistisch signifikant. Die Heterogenität zwischen den 3 Studien entsteht durch einen signifikanten Unterschied zwischen Garshick et al. [15] und Steenland et al. [39], ΔKonst1 : $p = 0,021$. Danach weicht die Garshick-Studie signifikant im Ausgangsniveau des Risikos von den anderen beiden Studien nach oben ab, was sich bereits in der Einzelanalyse andeutete (■ **Tab. 6**). Durch die Korrektur der Garshick-Koeffizienten sinkt die Heterogenität im Risikoniveau beträchtlich: $F(2, 6) = 0,50$ ($p = 0,63$). Dies lässt sich aus der Einzelanalyse nicht erkennen.

Tab. 9 Lineare Regression mit festen Effekten zu log RR mit Gewichten proportional zum Inversen der jeweiligen Varianz, ohne Adjustierung nach Studien

Log RR	Koef.	SE	t	p-Wert	95 % CI	
Dosis	0,00076	0,00026	2,97	0,018 (0,003)	0,00017 (0,00026)	0,00135 (0,00126)
Konstante	0,130	0,072	1,80	0,11	–0,0365	0,297
<i>Crump et al. [13], REC4</i>						
Dosis	0,00033	0,00022	1,46	0,18 (0,15)	–0,00019 (–0,00011)	0,00084 (0,00076)
Konstante	0,208	0,067	3,10	0,015	0,053	0,336
<i>Crump et al. [13], REC6</i>						
Dosis	0,00034	0,00020	1,74	0,12 (0,083)	–0,00011 (–0,00004)	0,00080 (0,00073)
Konstante	0,211	0,054	3,90	0,005	0,860	0,335

Koeffizient (Koef.) der kumulierten Exposition (Dosis) gegenüber DME in $\mu\text{g}/\text{m}^3$ -Jahre und Absolutterm (Konstante) mit Standardabweichung (SE); t-Teststatistik zum Koeffizienten mit zugehörigem p-Wert und 95 % Konfidenzintervall zum Koeffizienten (95 % CI). In Klammern die Ergebnisse bei Anwendung der Standardnormalverteilung. Eingeschlossen wird Silverman et al. [36], wie von Vermeulen et al. [44] verwendet, und entsprechend der Reanalyse von Crump et al. [13] mit Adjustierung für die Radonexposition (REC4, REC6; **Tab. 3**).

Tab. 10 Lineare Regression mit festen Effekten zu log RR mit Gewichten proportional zum Inversen der jeweiligen Varianz, mit Adjustierung nach Studien

Log RR	Koef.	SE	t	p-Wert	95 % CI	
Dosis	0,00106	0,00020	5,35	0,002 ($< 0,001$)	0,00057 (0,00067)	0,00154 (0,00144)
Δ Konst1	0,249	0,080	3,11	0,021	0,052	0,445
Δ Konst2	–0,035	0,126	–0,28	0,79	–0,343	0,272
Konstante	–0,059	0,079	–0,74	0,49	–0,252	0,135
<i>Crump et al. [13], REC4</i>						
Dosis	0,00053	0,00021	2,49	0,047 (0,013)	0,00001 (0,00011)	0,00105 (0,00094)
Δ Konst1	0,163	0,095	1,71	0,14	–0,070	0,396
Δ Konst2	–0,167	0,169	–0,98	0,36	–0,581	0,248
Konstante	0,091	0,091	1,01	0,35	–0,130	0,313
<i>Crump et al. [13], REC6</i>						
Dosis	0,00054	0,00017	3,12	0,021 (0,0018)	0,00012 (0,00020)	0,00096 (0,00087)
Δ Konst1	0,164	0,069	2,40	0,054	–0,003	0,332
Δ Konst2	0,041	0,113	–0,37	0,73	–0,319	0,236
Konstante	0,088	0,068	1,30	0,24	–0,079	0,255

Koeffizient (Koef.) der kumulierten Exposition (dose) gegenüber DME in $\mu\text{g}/\text{m}^3$ -Jahre und Absolutterm (Konstante) mit Standardabweichung (SE); t-Teststatistik zum Koeffizienten mit zugehörigem p-Wert und 95 % Konfidenzintervall zum Koeffizienten (95 % CI). In Klammern die Ergebnisse bei Anwendung der Standardnormalverteilung. Eingeschlossen wird Silverman et al. [36], wie von Vermeulen et al. [44] verwendet, und entsprechend der Reanalyse von Crump et al. [13] mit Adjustierung für die Radonexposition (REC4, REC6; **Tab. 3**).

Konstante: Steenland et al. [39] (Referenz-Studie).

Δ Konst1: Garshick et al. [15] – Steenland et al. [39].

Δ Konst2: Silverman et al. [36] – Steenland et al. [39].

Werden die Originalangaben nach Silverman et al. [36] durch die Ergebnisse von Crump et al. [13] ersetzt (REC4 oder REC6), so halbieren sich die Dosiskoeffizienten in der Meta-Regression.

Tab. 11 zeigt Ergebnisse, wenn eine gemischte lineare Regression zu den 3 Studien gefittet wird.

Tab. 11 ergibt eine sehr gute Übereinstimmung mit dem publizierten Befund in Vermeulen et al. [44] (Dosis = 0,00098, 95 % CI 0,00055–0,00141; Konstante = 0,088). Wird zusätzlich der Dosiskoeffizient als zufälliger Effekt geschätzt, ändern sich die Ergebnisse nicht. Ohne eine effektive Skalierung der Gewich-

te (Ergebnis: Dosiskoeffizient = 0,00093, Konstante = 0,096) weichen die Befunde etwas von Vermeulen et al. [44] ab. Wird Garshick et al. [15] (modifiziert) in die Metaanalyse eingeschlossen, sinkt der geschätzte Meta-Risikokoeffizient und das zugehörige Signifikanzniveau (Teststatistik ändert sich von $Z = 6,74$ auf $Z = 4,58$), allerdings ist der positive Dosis-Wirkungs-Zusammenhang stets signifikant (**Tab. 11**).

Wie in den Modellen mit festen Effekten (**Tab. 10**) halbieren sich in der gemischten Regression (**Tab. 11**) die Dosiskoeffizienten in der Meta-Regression, wenn die Originalangaben nach Silverman et al. [36] durch die Befunde von Crump et al. [13] (REC4 oder REC6) ersetzt werden.

Crump [10] führte ebenfalls eine Nachberechnung mit einem gemischten Modell durch, sah dabei aber keine Möglichkeit, die notwendigen Daten zu erhalten, um die Greenland/Longnecker-Methode einzusetzen (vgl. übereinstimmend unsere Erläuterungen im Methodenkapitel „Reproduktion der Ergebnisse und Variation der Modellierung“). Crump schreibt: „I (...) reran the analysis of Vermeulen et al. (...), except that I did not model the dependence among the ORs from the same study. (I did not have access to data needed to model that dependence.) My analysis yielded a regression parameter [0.88; 95 % confidence interval (CI): 0.65, 1.11] similar to that obtained by Vermeulen et al. (0.98; 95 % CI: 0.55, 1.41)“. Crump [10] wählte offenkundig eine abweichende Einheit der DME-Exposition von $1000 \mu\text{g}/\text{m}^3$ -Jahre. Der von ihm berichtete Dosiskoeffizient (0,88) weicht stärker ab als das Ergebnis der hier durchgeführten Nachberechnung (0,97), das sehr gut zu den Vermeulen-Ergebnissen passt (0,98). Es ist unklar, ob Crump [10] eine Skalierung der Gewichte vornahm, denn ohne Skalierung weicht auch die Nachberechnung in diesem Bericht etwas nach unten ab (0,93), wenn auch nicht so ausgeprägt wie bei Crump 2014 [10].

Variation der Modellierung

Da für Steenland et al. [39] keine Angaben zu Personenjahren und Fallzahlen je Expositionskategorie vorlagen, kann die-

Tab. 11 Gemischte lineare Regression zu log RR mit einem zufälligen Achsenabschnitt mit Gewichten auf der ersten Ebene (Ebene der Expositions-kategorien) proportional zum Inversen der jeweiligen Varianz und mit den Summen dieser Gewichte als Studiengewichte auf der zweiten Ebene (Studienebene)

Log RR	Koef.	Robust SE	Z	p-Wert	95 % CI	
Dosis	0,00097	0,00014	6,74	<0,001	0,00069	0,00125
Konstante	0,084	0,111	0,75	0,45	–0,134	0,300
<i>Garshick (modifiziert)</i>						
Dosis	0,00087	0,00019	4,58	<0,001	0,00050	0,00124
Konstante	0,024	0,062	0,39	0,70	–0,097	0,145
<i>Crump et al. [13], REC4</i>						
Dosis	0,00045	0,00021	2,10	0,035	0,00003	0,00086
Konstante	0,178	0,094	1,90	0,057	–0,005	0,362
<i>Crump et al. [13], REC4 und Garshick (modifiziert)</i>						
Dosis	0,00048	0,00023	2,10	0,036	0,00003	0,00092
Konstante	0,097	0,040	2,39	0,017	0,018	0,176
<i>Crump et al. [13], REC6</i>						
Dosis	0,00053	0,00026	2,05	0,040	0,00002	0,00103
Konstante	0,172	0,101	1,71	0,088	–0,025	0,369
<i>Crump et al. [13], REC6 und Garshick (modifiziert)</i>						
Dosis	0,00052	0,00025	2,09	0,036	0,00003	0,00101
Konstante	0,095	0,051	1,87	0,061	–0,044	0,194

Skalierung der Gewichte der ersten Ebene effektiv auf die Gewichte der zweiten Ebene. Koeffizient (Koef.) der kumulierten Exposition („dose“) gegenüber DME in $\mu\text{g}/\text{m}^3\text{-Jahre}$ und Absolutterm (Konstante) mit Standardabweichung (SE); Z-Teststatistik zum Koeffizienten mit zugehörigem P-Wert und 95 %-Konfidenzintervall zum Koeffizienten, 95 % CI. In Klammern die Ergebnisse bei Anwendung der Standardnormalverteilung. Eingeschlossen wird Silverman et al. [36], wie von Vermeulen et al. [44] benutzt, und entsprechend der Reanalyse von Crump et al. [13] mit Adjustierung für die Radonexposition (REC4, REC6, vgl. Tab. 3). Eingeschlossen wird Garshick et al. [15], wie von Vermeulen et al. [44] benutzt, bzw. Garshick et al. [15] (modifiziert).

se Arbeit nicht in Auswertungen mit der Greenland/Longnecker-Methode aufgenommen werden. Gemeinsame Analysen zu Garshick et al. [15], Silverman et al. [36] und Möhner et al. [21] (adaptiert) konnten jedoch erfolgen. Die Ergebnisse sind in Tab. 12 dargestellt.

Tab. 12 zeigt, dass die linearen Regressionen mit festen Effekten zu log RR mit Gewichten proportional zum Inversen der jeweiligen Varianz und mit Adjustierung nach Studien leicht ausgeprägtere Wirkungskoeffizienten und etwas niedrigere p-Werte als Modelle mit zufälligen Effekten ergeben (dies stimmt qualitativ mit den Befunden überein, die im Ergebniskapitel „Gemeinsame Analyse der Studien“ dargestellt sind). Wird allerdings auch der Dosiskoeffizient als zufälliger Effekt ermittelt, ergibt sich ein ähnlicher Wert (Koeffizient = 0,00059, $p = 0,17$) zum Modell mit festen Effekten. Die Analysen mit der von Methodikern empfohlenen und die internen Korrelationen berücksichtigenden Greenland/Longne-

cker-Methode ergaben dagegen einen geringeren Koeffizienten als in dem Modell mit festen Effekten, mit einem niedrigeren p-Wert = 0,035 (signifikant). Der Dosiskoeffizient liegt in dieser gemeinsamen Analyse von Garshick et al. [15] (modifizierte Koeffizienten), Silverman et al. [36] und Möhner et al. [21] (adaptiert) bei Anwendung der Greenland/Longnecker-Methode (0,00032) um etwa einen Faktor 3 unterhalb der Schätzung in der Hauptanalyse von Vermeulen et al. [44] (0,00098).

Die von Vermeulen et al. [44] berichteten Ergebnisse stimmen weitgehend mit den Schätzungen der Modelle mit zufälligen Effekten überein (vgl. Ergebniskapitel zu „Gemeinsame Analyse zu den Studien“). Die Methode von Greenland/Longnecker kann aber ohne den Zugang zu weiteren Daten hier nicht eingesetzt werden. Vermeulen und Mitarbeiter erhielten offensichtlich Zugang zu diesen Daten aus Steenland et al. [39], haben sie in ihrer Arbeit leider auch im Supplement nicht offen gelegt.

Werden die Originalangaben nach Silverman et al. [36] durch die Ergebnisse von Crump et al. [13] ersetzt (REC4 oder REC6), so finden sich qualitativ dieselben Aussagen zu den 3 Analyseverfahren. Übereinstimmend halbieren sich die Dosiskoeffizienten in der Meta-Regression durch die Datenvariation. Die p-Werte stimmen trotz der unterschiedlichen Modellierungen weitgehend überein. Eine Ausnahme ergibt sich für die Regression mit zufälligem Achsenabschnitt zu REC6, wo der Wert niedriger ausfällt ($p = 0,25$). Dies deutet, wie beim abweichenden Ergebnis der Greenland/Longnecker-Regression bei Auswertung mit den Silverman-Originaldaten ($p = 0,035$), auf Instabilitäten in der Varianzschätzung hin.

Um Diskussionen um eine mögliche Unterschätzung der Wirkung zu vermeiden, werden alle weiteren Auswertungen mit adjustierten Regressionsmodellen mit festen Effekten durchgeführt. In Tab. 11 und 12 stimmen die Dosiskoeffizientenschätzungen (feste Effekte nach Studien adjustiert und zufällige Effekte) fast überein und sind zudem beinahe identisch mit den Angaben in Vermeulen et al. [44]. Ein Vergleich der Intervallschätzungen und p-Werte weist auf eine höhere Stabilität der Regressionsmodelle mit festen Effekten hin. Somit ist eine Festlegung auf adjustierte Regressionsmodelle mit festen Effekten im Sinne einer Hauptanalyse (zentraler Modellierungsansatz) für dieses Forschungsprojekt gerechtfertigt.

Einfluss der gewählten Inputdaten

Eine gleichzeitige Analyse der 3 Studien von Garshick et al. [15] (modifiziert), Silverman et al. [36] und Möhner et al. [21] (adaptiert) ergibt unabhängig von der gewählten Methodik einen deutlich niedrigeren Wirkungskoeffizienten in Tab. 12 als in der Hauptanalyse nach Vermeulen et al. 2014.

Beispiel: Für die Regressionsmodelle mit festen Effekten und Adjustierung für Studien ergab sich eine Reduktion im Wirkungskoeffizienten um etwa 50 %, denn die reproduzierte Hauptanalyse nach Vermeulen et al. [44] ergab einen Koeffizienten von 0,0011 (pro $1 \mu\text{g}/\text{m}^3\text{-Jahre}$), $p = 0,002$, und ein 95 % CI von 0,00057, bis 0,00154 (Tab. 10).

Tab. 12 Gemeinsame Analyse von Garshick et al. [15] (modifizierte Koeffizienten), Silverman et al. [36] und Möhner et al. [21] (adaptiert)

	Koef.	p-Wert	95 % CI	
Feste Effekte, adjustiert	0,00054	0,14 (0,083)	– 0,00026 (– 0,00007)	0,00133 (0,00115)
Zufälliger Achsenabschnitt, skaliert	0,00024	0,36	– 0,00027	0,00074
Greenland/Longnecker-Methode	0,00032	0,035	0,00002	0,00062
<i>Crump et al. [13], REC4</i>				
Feste Effekte, adjustiert	0,00020	0,35 (0,30)	– 0,00029 (– 0,00018)	0,00069 (0,00058)
Zufälliger Achsenabschnitt, skaliert	0,00010	0,11	– 0,00002	0,00022
Greenland/Longnecker-Methode	0,00015	0,31	– 0,00014	0,00044
<i>Crump et al. [13], REC6</i>				
Feste Effekte, adjustiert	0,00013	0,33 (0,28)	– 0,00018 (– 0,00010)	0,00042 (0,00035)
Zufälliger Achsenabschnitt, skaliert	0,000073	0,025	0,00001	0,00014
Greenland/Longnecker-Methode	0,00012	0,44	– 0,00019	0,00043

Vergleich der Schätzung der Wirkung für die kumulierte Exposition auf log RR, ermittelt mit unterschiedlichen Regressionsverfahren. Koeffizient (Koef.) der kumulierten Exposition (Dosis) gegenüber DME in $\mu\text{g}/\text{m}^3$ -Jahre, zugehöriger p-Wert und 95 % Konfidenzintervall zum Koeffizienten (95 % CI). In Klammern die Ergebnisse bei Anwendung der Standardnormalverteilung im Modell mit festen Effekten. Eingeschlossen wird Silverman et al. [36], wie von Vermeulen et al. [44] verwendet, und entsprechend der Reanalyse von Crump et al. [13] mit Adjustierung für die Radonexposition (REC4, REC6; **Tab. 3**).

Als weiterer interessanter Befund ist festzuhalten, dass eine gleichzeitige Analyse der 3 Studien von Garshick et al. [15] (modifiziert), Silverman et al. [36] und Möhner et al. [21] (adaptiert) nur bei Anwendung der Greenland/Longnecker-Methode eine signifikante Dosis-Wirkungs-Beziehung findet (**Tab. 12**). Dies könnte an Instabilitäten in der Schätzung bei diesem komplexen Verfahren liegen.

Wird die Hauptanalyse nach Vermeulen et al. [44] mit festen Effekten und Adjustierung für Studien und mit modifizierten Garshick-Koeffizienten durchgeführt, so ergibt sich ein Koeffizient von 0,00098 (pro 1 $\mu\text{g}/\text{m}^3$ -Jahre), $p = 0,006$, und ein 95 % CI von 0,00041 bis 0,00155, etwas ausgeprägter aber ähnlich zu den Ergebnissen der Regressionsrechnung mit zufälligen Effekten (**Tab. 11**).

Somit bewirkt die Modifikation der Garshick-Koeffizienten lediglich eine Reduktion im Meta-Koeffizienten um etwa 7 % und die statistische Signifikanz bleibt erhalten, obgleich der p-Wert von 0,002 auf 0,006 ansteigt. Durch die Korrektur der Garshick-Koeffizienten sinkt

die Heterogenität im Ausgangsniveau des Risikos beträchtlich: $F(2,6) = 0,50$, $p = 0,63$, aber ohne Korrektur: $F(2,6) = 5,5$, $p = 0,044$.

Tab. 10 bis 12 zeigen übereinstimmend, dass eine Ersetzung der Originaldaten von Silverman et al. [36] durch die Ergebnisse von Crump et al. [13] (REC4 oder REC6) zu einer ungefähren Halbierung der Dosiskoeffizienten in den Meta-Regressionen führt. Werden Steenland et al. [39], Crump et al. [13] und Garshick et al. [15] (modifiziert) gemeinsam analysiert, so ist der Dosis-Wirkungs-Zusammenhang in adjustierten Regressionsmodellen mit festen Effekten und bei Bewertung mit der Standardnormalverteilung signifikant (REC4: $p = 0,03$, REC6: $p = 0,01$), aber deutlich schwächer als in einer Analyse der von Vermeulen et al. [44] verwendeten Originaldaten ($p < 0,000001$). Entsprechend verschiedenen fallen die unteren Konfidenzintervallgrenzen zum Dosiskoeffizienten aus: 0,00005 und 0,00010 bzw. 0,00067. Hier finden sich Unterschiede um einen Faktor von mindestens 12 oder 6.

Abb. 2 gibt einen Überblick zu den Ergebnissen der Meta-Regressionen mit festen Effekten bei Variation der Inputdaten und Annahme einer Normalverteilung zur statistischen Bewertung. Modelle mit festen Effekten werden dargestellt, um die Koeffizienten nicht zu niedrig einzuschätzen. Die Normalverteilung wurde gewählt, um eine Überschätzung der p-Werte auszuschließen (d. h. die dargestellten Konfidenzintervalle sind keinesfalls zu weit). Abweichend von Vermeulen et al. [44] wird als Expositionseinheit mg/m^3 -Jahre = $1000 \times \mu\text{g}/\text{m}^3$ -Jahre gewählt, um die Darstellung übersichtlich zu halten. Die Replikation der Analyse, wie in Vermeulen et al. [44] durchgeführt, d. h. Steen_Silv_Garsh (mit Adj), ergibt den höchsten Risikoschätzer unter allen Datenvariationen. Werden die Modifikationen der Silverman-Fall-Kontroll-Studie nach Crump et al. [13] (REC4, REC6) und zu Garshick et al. [15] in die Analyse eingeschlossen, liegen die Risikoschätzer deutlich niedriger (Steen_SilvREC4_Garsh_mod, Steen_SilvREC6_Garsh_mod: Die Koeffizienten sind ungefähr halbiert gegenüber der Vermeulen-Analyse). Bei Ausschluss der Steenland-Arbeit, die nach Ansicht der Autoren nur explorativen Charakter hat, ergeben die modifizierten Risikoschätzer bei gleichzeitiger Berücksichtigung von Möhner et al. [21] (adaptiert) in der Metaanalyse keine signifikante Risikoerhöhung (SilvREC4_Garsh_mod_Moehn, SilvREC6_Garsh_mod_Moehn). Der Meta-Koeffizient beträgt etwa 10–20 % des Wertes der bei Vermeulen et al. [44] in der Hauptanalyse gefunden wurde.

Nichtlinearität: Schwellenwertsuche

Die Hauptanalyse von Vermeulen et al. [44] wurde wiederholt (Verfahren: adjustierte Regressionsmodelle, feste Effekte), aber unter Berücksichtigung eines möglichen Schwellenwertes. Wird die Auswertung mit korrigierten Koeffizienten in Garshick-Koeffizienten durchgeführt (**Tab. 2: [15]**, modifiziert), so ermittelt die Analyse einen Schwellenwert bei 150 $\mu\text{g}/\text{m}^3$. Die Schwelle ist aber nicht statistisch signifikant von Null verschieden.

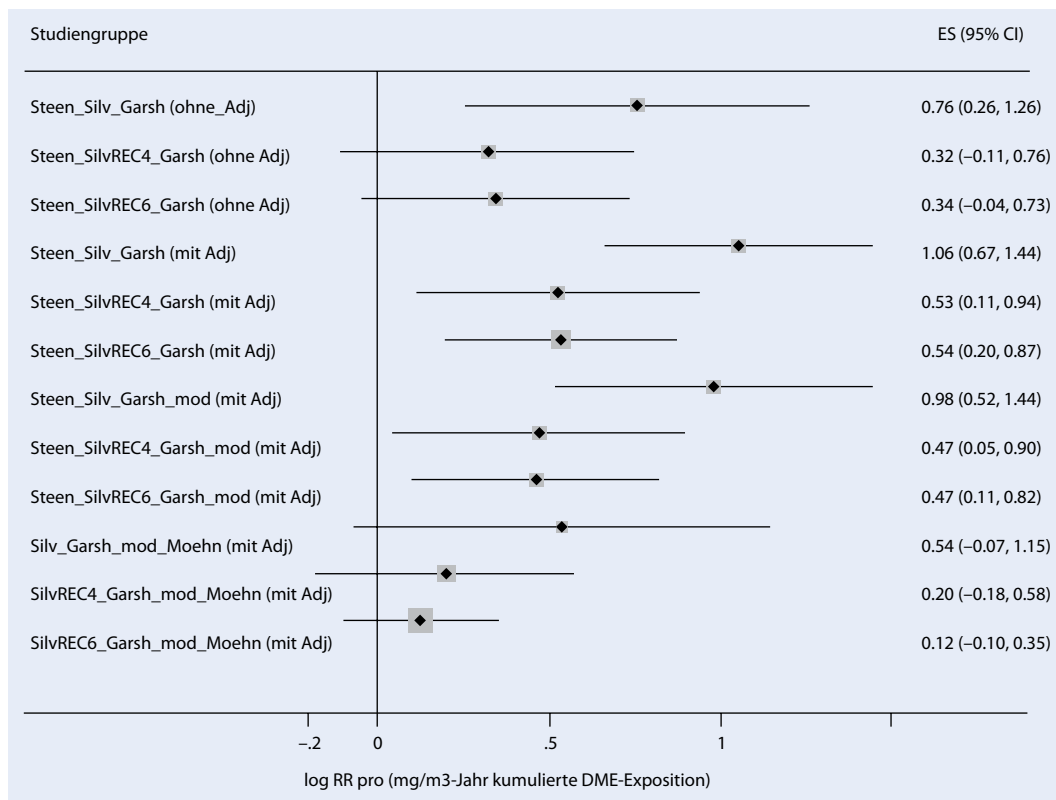


Abb. 2 ▲ Metaanalysen. Lineare Regression mit *festen* Effekten zu log RR mit Gewichten proportional zum Inversen der jeweiligen Varianz (die Boxgröße veranschaulicht die Präzision der Meta-Schätzung). Koeffizienten („effect size“, ES) der kumulierten Exposition (Dosis) gegenüber DME in mg/m³-Jahre mit 95 % Konfidenzintervall zum Koeffizienten, berechnet mit Hilfe der Standardnormalverteilung. Ausgewertet werden Steenland et al. [39] und Silverman et al. [36], wie von Vermeulen et al. [44] verwendet, und entsprechend der Reanalyse von Crump et al. [13] mit Adjustierung für die Radonexposition (REC4, REC6; ■ Tab. 3), Garshick et al. [15], wie von Vermeulen et al. [44] verwendet, und Garshick et al. [15] (modifiziert) sowie die Daten nach Möhner et al. [21] (adaptiert)

Wird Silverman et al. [36] (■ Tab. 1) durch die Ergebnisse aus Crump et al. [13] (■ Tab. 4) ersetzt, so zeigt die Analyse ebenfalls einen Schwellenwert. Allerdings ist dieser Schätzer aufgrund der deutlich schwächeren Expositions-Wirkungs-Beziehung statistisch schlechter zu fassen und zeigt eine weite Unsicherheit (Schwelle bei 90 µg/m³-Jahre, 95 % CI 0 µg/m³-Jahre bis 361 µg/m³-Jahre).

Diskussion

Vermeulen et al. [44] publizierten eine Meta-Regressionsanalyse zu 3 wesentlichen epidemiologischen US-amerikanischen Studien [15, 36, 39], die den Zusammenhang zwischen Dieselmotor-emissionen (DME, engl. DEE), gemessen in der kumulierten Exposition gegenüber elementarem Kohlenstoff (EC) in µg/m³-Jahre und der Lungenkrebsmortalität darstellten. Die Autoren beschrieben in

ihrer Metaanalyse eine statistisch signifikante Dosis-Wirkungs-Beziehung und erhöhte Krebsrisiken selbst bei sehr niedrigen Belastungen. Die vorliegende Reanalyse konnte die Einzel- und Hauptbefunde weitgehend aus den publizierten Studiendaten reproduzieren. Unter allen von uns durchgeführten Metaanalysen ergab die Auswertung der Daten, wie von Vermeulen et al. [44] verwendet, den höchsten Risikoschätzer. Allerdings ergab eine Untersuchung zur Heterogenität im Ausgangsniveau des Risikos – von Vermeulen et al. [44] wird hierzu nicht berichtet – ausgeprägte Unterschiede zwischen den 3 Studien (signifikant auf dem 5 %-Niveau). Alle 3 Studien sollten einheitlich ein Ausgangsniveau von RR = 1 bei einer kumulierten DME-Exposition von 0 µg/m³-Jahre zeigen. Eine gemeinsame Analyse der 3 Studien ist somit aus statistischer Sicht kritisch zu bewerten [1]. Eine Zusammenführung von Studien wird in an-

deren Situationen bei selbst deutlich geringer ausgeprägter Heterogenität abgelehnt [8, 26]. Diese Unsicherheit im Ausgangsniveau problematisiert die Verwendung der Analyse von Vermeulen et al. [44] für Risikoabschätzungen in niedrigeren Expositionsbereichen. Die folgende Aussage von Vermeulen et al. [44] ist insofern zu relativieren: „Formal tests of heterogeneity of estimates among the studies were of limited value due to the small number of data points for each study.“

Eine Korrektur der Koeffizienten in Garshick et al. [15] (vgl. [24]) bewirkte eine beträchtliche Reduktion der Heterogenität im Ausgangsniveau zwischen den 3 Studien ($p=0,63$). Diese Korrektur war indiziert, da die von Vermeulen et al. [44] verwendeten Koeffizienten fehladjustiert sind [24]: Die kumulierte Exposition wurde zusätzlich nach der in ihr bereits enthaltenen Expositionsdauer adjustiert, so dass der Risikokoeffizient nicht die Wir-

kung der kumulierten Exposition schätzt. Die Modifikation ergab deutlich niedrigere Risikoschätzungen als zu dieser Studie von Vermeulen et al. [44] berichtet wird: Der korrigierte Dosiskoeffizient ist mehr als eine 10er-Potenz kleiner und weit entfernt von einer Signifikanz ($p=0,6$). Die Korrektur der Garshick-Koeffizienten bewirkt allerdings lediglich eine Reduktion im Meta-Koeffizienten um etwa 7 % und die statistische Signifikanz bleibt erhalten, obgleich der p-Wert von 0,002 auf 0,006 ansteigt.

Zur korrekten Bewertung der Garshick-Studie ist ein weiterer Aspekt wichtig (vgl. den Leserbrief von [24], inklusive Autorenantwort). Garshick et al. [15] haben doppelt für das Geburtsjahr adjustiert, wenn auch mit unterschiedlicher Feinheit, so dass die Modelle nicht wegen Kollinearität zusammenbrechen. Allerdings kann eine solche Überadjustierung die Koeffizientenschätzung verzerren. Die Autoren antworteten, dass sie diese Doppelberücksichtigung durchführten, um sinnvolle Ergebnisse zu erhalten. Erst dann sei die sog. Proportional Hazards Assumption erfüllt gewesen. Dieser Erläuterung ändert jedoch an der Doppelverwendung der Geburtsjahrsinformation und der dadurch verursachten potenziellen Überadjustierung nichts.

Der höchste in die Metastudie eingehende Expositionswert stammt aus der DEMS-Fall-Kontroll-Studie von Silverman et al. [36]: $1036 \mu\text{g}/\text{m}^3\text{-Jahre}$ mit einem $\text{OR}=2,83$ (95 % CI 1,28–6,26). Diese US Bergbau-Studie hat somit für die Metaanalyse besondere Relevanz. Zu den DEMS-Publikationen [3, 36] wurden kritische Anmerkungen und eine Liste offener Fragen veröffentlicht [25], auf die die Autoren mit einem Leserbrief reagierten [35]. Allerdings blieben viele Aspekte offen (siehe die Autorenantwort von Morfeld in [35]). Diese offenen Fragen ließen sich z. T. durch Zusatzanalysen beantworten, die aber einen Zugang zu den Originaldaten voraussetzen. Ein solcher Zugriff auf die Originaldaten der DEMS-Studie wurde bislang nur wenigen Forschern ermöglicht (Arbeitsgruppen um S. Moolgavkar und K. Crump), wobei derzeit unklar bleibt, ob diese Autoren uneingeschränkten Zugang zu allen Originaldaten hatten. Die Ergebnisse zur DEMS-Reana-

lyse wurden in dem HEI (Health Effects Institute) Webinar präsentiert [12, 21] und nach Peer-review in „Risk Analysis“ zur Publikation angenommen [13, 22].

In einem Leserbrief zu Vermeulen et al. [44] berichtete Crump [10], dass eine Auswertung der DEMS-Fall-Kontroll-Studie mit einem lag der Exposition um 5 Jahre zu deutlich niedrigeren Meta-Risikokoeffizienten führt als mit dem Originalwert von 15 Jahren. Crump weist zudem darauf hin, dass die anderen beiden eingeschlossenen Studien [15, 39] einen lag von 5 Jahren verwendeten. Die Autoren bestätigen qualitativ das Crump-Ergebnis, halten es aber nicht für wesentlich, da die Anpassung in der DEMS-Studie bei Verwendung eines „lags“ von 15 Jahren besser war und unvermeidlich Unterschiede in den Expositionserhebungen zwischen Studien bestehen. Vermeulen et al. [44] verwiesen aber auf die von Ihnen durchgeführten Sensitivitätsanalysen, in denen sie explorieren, welchen Einfluss ein unterschiedliches lagging der Exposition hat. Der Hinweis von Crump [10] zeigt allerdings weitere Unsicherheiten der Vermeulen-Metaanalyse auf, die ohne einen Zugriff auf die DEMS-Originaldaten nicht ermittelbar sind.

Moolgavkar et al. [23] führten eine Reanalyse der DEMS-Kohortenstudie von Attfield et al. [3] durch und weisen auf zwei wichtige Umstände hin:

1. Zeitabhängige Faktoren überlagern das Risikogeschehen, so dass Modellkoeffizienten nicht ohne Wechselwirkung mit dem Alter geschätzt werden sollten. Die Angabe eines isolierten Risikokoeffizienten – wie in Attfield et al. [3] beschrieben – macht nach Moolgavkar et al. [23] somit keinen Sinn.
2. Ein Bergwerk („limestone mine“, Kalksteinmine) stellt einen Ausreißer in den Daten dar. Die DME-Expositionen liegen in diesem Bergwerk am niedrigsten, aber die Risikoschätzungen sind die höchsten unter allen Bergwerken. Eine Cox-Regressions-Analyse der Daten zeigt (ohne gewisse Expositionshöhen vorweg auszuschließen), dass sich allein in der limestone mine ein signifikanter Expositions-Wirkungs-Zusammenhang ergibt ($p=0,0014$), der sich aber

in keinem anderen Bergwerk finden lässt. Wird die „limestone mine“ bei den Analysen ausgeschlossen, so steigt der p-Wert im Gesamtkollektiv von 0,02 auf 0,18 [22], d. h. es ergibt sich nach Ausschluss der „limestone mine“ keinen signifikanter Dosis-Wirkungs-Zusammenhang zwischen DME-Exposition und Lungenkrebssterblichkeit in der DEMS-Kohortenstudie. Obwohl diese Sonderrolle der „limestone mine“ in der Fall-Kontroll-Reanalyse von Crump et al. [13] nicht bestätigt werden konnte, stellt diese Sonderstellung ein erhebliches Interpretationsproblem für die Kohortenstudie und die Publikation von Attfield et al. [3] dar.

Crump et al. [13] reanalysierten die DEMS-Fall-Kontroll-Studie von Silverman et al. [36] und untersuchten zusätzlich den Einfluss von Kovariablen, die Silverman et al. [36] nicht in ihre finalen Modelle übernommen hatten. Die Analyse des Zusammenhangs von kumulierter DME-Exposition (lag = 15 Jahre) und Lungenkrebsmortalität ergab in der Originalarbeit einen Trend-p-Wert 0,001 (Tab. 3, [36]), den Crump und Mitarbeiter bestätigen konnten: $p=0,0006$ in der Reanalyse. Crump et al. [13] (Table III) berichteten einen Trend-p-Wert von $p=0,02$, wenn zusätzlich nach der Radonexposition adjustiert wird. Dies belegt, dass die Radonexposition einen deutlichen Einfluss auf den Zusammenhang besitzt und sich nach Berücksichtigung dieser Kovariablen die statistische Signifikanz der DME-Variablen abschwächt. Crump et al. [13] entwickelten zudem 6 neue DME-Expositionsmaße, alternativ zu den Schätzungen, die in Attfield et al. [3] und Silverman et al. [36] eingesetzt wurden. Werden diese alternativen DME-Expositionsmaße verwendet und gleichzeitig nach der Radonexposition adjustiert, so ergibt sich in keiner Konstellation ein signifikanter Zusammenhang zwischen kumulierter DME-Exposition und Lungenkrebsmortalität ($p \geq 0,17$ oder der Trend ist negativ). Auch für die Expositionsmaße, wie sie in Attfield et al. [3] und Silverman et al. [36] verwendet wurden, zeigen die Analysen von Crump et al. [13] keinen signifikanten Zusammenhang auf, sobald nicht

die Durchschnittswerte der Gruppen sondern die Individualdaten in der Analyse verwendet und nach Radonexpositionen adjustiert wird: $p \geq 0,65$ oder der Trend ist negativ. Die Autoren schreiben: „Most importantly, we used the radon concentration data for the DEMS cohort provided by the DEMS investigators. When adjustment was made for radon, a known human lung carcinogen, the effect of REC on the association with lung cancer mortality was confined to only the three DEMS REC estimates. Most notably, there was no evidence of an association with any of the six alternate REC estimates, including REC6. When T2 trend tests were conducted, based on the use of individual worker REC estimates, the results were less statistically significant and in many cases the trends were negative. Indeed, for miners who always worked underground, five of the six REC metrics exhibited negative trends.“

Moolgavkar et al. [23] und Crump et al. [13] folgerten, dass die DEMS-Auswertungen von Attfield et al. [3] und Silverman et al. [36] nicht geeignet sind, um allein in einer quantitativen Risikoanalyse verwendet zu werden. Auch für Metaanalysen und Grenzwertableitungen sind die Limitationen dieser DEMS-Daten zu berücksichtigen.

Wird die Arbeit von Crump et al. [13] (REC4, REC6) an Stelle von Silverman et al. [36] berücksichtigt, so ergibt sich in der hier durchgeführten Meta-Reanalyse von Vermeulen et al. [44] ein deutlich abgeschwächter Zusammenhang zwischen DME-Exposition und Lungenkrebssterblichkeit. ■ **Tab. 10 bis 12** zeigen übereinstimmend, dass eine entsprechende Ersetzung der Originaldaten zu einer ungefähren Halbierung der Dosiskoeffizienten in den Meta-Regressionen führt. Werden Steenland et al. [39], Crump et al. [13] und Garshick et al. [15] (modifiziert) gemeinsam analysiert, so ist der Dosis-Wirkungs-Zusammenhang in einem adjustierten Regressionsmodelle mit festen Effekten eben noch signifikant ($p = 0,03$), also deutlich unsicherer als in der Originaldatenanalyse ($p < 0,000001$).

Die empfohlene Greenland/Longnecker-Technik konnte nicht auf die älteste der Studien [39] angewendet werden, da hierzu wichtige Zusatzinformationen in

der Publikation fehlen. Mit dieser statistischen Methode konnten allerdings Metaanalysen zu Garshick et al. [15], Silverman et al. [36] und Möhner et al. [21] durchgeführt werden. Eine gemeinsame Analyse von Garshick et al. [15] (modifiziert), Silverman et al. [36] und Möhner et al. [21] (adaptiert) ergab unabhängig von der gewählten Methodik ähnliche Koeffizientenschätzungen, unabhängig vom Auswertungsverfahren (zufällig, fest, Greenland/Longnecker), wobei die Greenland/Longnecker-Schätzung zwischen den Ergebnissen aus dem Modell mit zufälligen Effekten und dem adjustierten Modell mit festen Effekten lag. Die Fokussierung der Analysen auf das adjustierte Modell mit festen Effekten, mit dem in allen Datensituationen Auswertungen möglich waren, ist somit gerechtfertigt. Der niedrigste Risikoschätzer ergab sich nach Ausschluss der explorativen Studie von Steenland et al. [39] in einer Analyse der 3 Studien Garshick et al. [15] (modifiziert), Silverman et al. [36] (mod. nach [13]) und Möhner et al. [21] (adaptiert). Der Meta-Koeffizient sank in dieser Auswertung auf etwa 10–20 % des Wertes, der von Vermeulen et al. [44] als Ergebnis ihrer Hauptanalyse publiziert wurde. Des Weiteren verliert der Zusammenhang zwischen DME-Exposition und Lungenkrebsmortalität in dieser Analyse die statistische Signifikanz.

Vermeulen et al. [44] führten aus: „We were not able to investigate other model forms in our meta-regression, beyond the linear and spline curves because of the limited number of data points. If nonlinear exposure-response curves were actually a better fit (e. g., attenuation at higher exposures, for which there is some evidence in Silverman et al. [36], then this might change the estimate burden of disease due to DEE.“ Wird die Metaanalyse zu den 3 US-amerikanischen Studien [15, 36, 39] um eine Suche nach Schwellen erweitert, also die Dosis-Respons-Beziehung auf Nichtlinearität untersucht, so ergibt sich mit korrigierten Garshick-Koeffizienten ein Schwellenwert für die kumulierte DME-Exposition bei $150 \mu\text{g}/\text{m}^3$ -Jahre. Dieser ist allerdings nicht signifikant von Null verschieden.

Sun et al. [41] erstellten einen Überblick zu den Ergebnissen von 42 Kohorten- und 32 Fall-Kontroll-Studien zur Fra-

ge des Zusammenhangs zwischen DME-Expositionen und Lungenkrebs. Die Autoren folgerten: „Overall, neither cohort nor case-control studies indicate a clear exposure-response relationship between DE exposure and lung cancer. Epidemiological studies published to date do not allow a valid quantification of the association between DE and lung cancer.“ Obgleich diese Arbeit nicht das methodische Niveau des metaanalytischen Zugangs von Vermeulen et al. [44] erreicht, so belegen die variierenden Studienergebnisse in Sun et al. [41] die Unsicherheit, die den epidemiologischen Studien zur Zusammenhangsfrage DME und Lungenkrebs zu Grunde liegt.

Die Reanalyse der deutschen Kalibergbaustudie [29] durch Möhner et al. [21] führte in ähnlicher Weise, wie in unserer Analyse der Vermeulenarbeit, zu einer klar abgeschwächten und anderen Aussage im Vergleich zur Originalarbeit: „Only for very high cumulative dose, corresponding to at least 20 years of exposure in the production area, some weak hints for a possible risk increase could be detected.“

Die hier durchgeführte Meta-Regressionsanalyse hat deutliche Beschränkungen. Wesentliche Limitationen ergeben sich aus dem Umstand, dass weder DME-Konzentrationswerte (nur Angaben zu kumulierten Expositionen) noch die Individualdaten für dieses Forschungsprojekt zur Verfügung standen. Schwellenwertanalysen für Stäube sollten sich auf eine Konzentrationsschwelle konzentrieren (vgl. Diskussion in [27]). Empirische Erkenntnisse zu Schwellen einer Quarzstaubeinwirkung, ermittelt an der deutschen Porzellanarbeiterkohorte, ergaben für die kumulierte Exposition keine Schwelle, jedoch für die Konzentration [27]. Ähnlich muss in diesem Fall vermutet werden, dass durch Beschränkung der Daten auf Angaben zur kumulierten DME-Exposition von einer Unterschätzung des tatsächlichen Konzentrations-schwellenwertes auszugehen ist. Die statistische Signifikanz des Befundes würde deutlich klarer ausfallen, wären die Originalstudien unter Berücksichtigung von Konzentrationswerten analysiert worden und in dieser Form in die Metaanalyse eingegangen. Die Daten gehen allein in aggregierter Form (Gruppendaten) in die

Analyse ein (■ Tab. 1), während die Originaldaten individuell sind. Solche Analysen zu vorab schon verdichteten Daten sind problematisch, da sich je nach Art der Rasterung andere Ergebnisse einstellen können und die Präzision durch die Kategorisierung abnimmt. Grundsätzlich führen Kategorisierungen zu Informationsverlust, potenziellen Verzerrungen und „Power“-Einbußen [34].

Diese genannten Limitationen betreffen vor allem auch die Arbeit von Vermeulen et al. [44]. Obgleich alle Hauptautoren der Einzelstudien zugleich Koautoren der Metaanalyse sind, haben Vermeulen et al. [44] darauf verzichtet, die Originaldaten zusammenzuführen und zu analysieren, sondern sich nur auf die gruppierten Daten aus den Ergebnistabellen der Publikationen beschränkt. Belastbare Analysen sollten sich jedoch auf Individualdaten beziehen und die DME-Konzentration als wesentliche Variable berücksichtigen.

Die hier durchgeführten Meta-Regressionen zeigen erhebliche Variationen der Ergebnisse, je nach eingeschlossenen Studiendaten oder gewählten Analyseverfahren. Die von Vermeulen und Mitarbeitern durchgeführte Datenauswahl führte zu den höchsten Risikoschätzern in unserer Metaanalyse (statistisch signifikant). Der niedrigste Risikoschätzer ergab sich nach Ausschluss der explorativen Studie von Steenland et al. [39] in einer Analyse der 3 Studien Garshick et al. [15] (modifiziert), Silverman et al. [36] (modifiziert nach [13]) und Möhner et al. [21] (adaptiert). Der Meta-Koeffizient sank in dieser Auswertung auf ca. 10–20 % des Wertes, der von Vermeulen et al. [44] als Ergebnis ihrer Hauptanalyse publiziert wurde. Des Weiteren verliert der Zusammenhang zwischen DME-Exposition und Lungenkrebsmortalität in dieser Analyse die statistische Signifikanz. Die aus Vermeulen et al. [44] abgeleiteten Risikoschätzungen im sehr niedrigen Expositionsbereich und zugehörige Grenzwertableitungen [32] sind daher nicht überzeugend, nicht zuletzt da die Daten – nach Korrektur der Garshick-Koeffizienten – auch grundsätzlich mit einem Schwellenwert verträglich sind.

Toxikologische Ergebnisse der aktuellen ACES-Studie aus kontrollierten Lang-

zeitversuchen mit Ratten zeigen bei lebenslanger Exposition mit „New Technology Diesel Exhaust“ (NTDE) gegenüber früheren Studien mit *traditionellen* DME (TDE) aus Dieselmotoren ohne Partikelminderung und ohne sonstige Abgasnachbehandlung u. a. weder eine Tumorinduktion noch Präkanzerosen [17]. In der Zusammenfassung der IARC-Workshopergebnisse zur Neubewertung von DME in 2012 wurde darauf hingewiesen, dass nur toxikologische Daten mit Motoren- und Kraftstofftechnologien bis zum Jahr 2000 berücksichtigt wurden [4]. Damit bezieht sich die IARC-Einstufung aus 2012 auch unter Nutzung der epidemiologischen Daten von Steenland et al. [39], Garshick et al. [15] und Silverman et al. [36] ausschließlich auf ein mögliches krebserzeugendes Potenzial von Dieselmotoremissionen ohne zeitgemäße Abgasnachbehandlung (TDE). Die neuen Tierversuchsergebnisse der ACES-Studie unterstreichen jedoch die Notwendigkeit der Berücksichtigung von Motoren- und Abgasnachbehandlungstechniken in der Diskussion um Arbeitsplatzgrenzwerte und mögliche Lungenkrebsrisiken durch DME. Auch McClellan et al. [20] weisen explizit auf die qualitativen und quantitativen Unterschiede zwischen TDE und NTDE hin und empfehlen, dass bei der Beurteilung kanzerogener Risiken diese Unterschiede berücksichtigt werden sollten.

Unabhängig von den genannten grundsätzlichen Beschränkungen hat die vorliegende Reanalyse also aufgezeigt, dass die Ergebnisse aus der Meta-Regressionsstudie von Vermeulen et al. [44] nicht in einer Risikoschätzung vorbehaltlos verwendet werden sollten, denn die Ergebnisse variieren ausgeprägt in Abhängigkeit von den gewählten Inputdaten und den verwendeten statistischen Verfahren. Insbesondere ist eine Anwendung der Ergebnisse von Vermeulen et al. [44] im Niedrigexpositions-bereich kontraindiziert.

Fazit

- Vermeulen et al. [44] publizierten eine Meta-Regressionsanalyse zu 3 wesentlichen epidemiologischen US-amerikanischen Studien zum Zusammenhang

zwischen Dieselmotoremissionen (DME) und Lungenkrebs. Sie ermittelten eine statistisch signifikante Dosis-Wirkungs-Beziehung und erhöhte Krebsrisiken selbst bei sehr niedrigen Belastungen mit DME.

- Die vorliegende Reanalyse kann die von Vermeulen et al. [44] berichteten Einzel- und Hauptbefunde weitgehend aus den publizierten Studiendaten reproduzieren.
- Die Meta-Regressionen zeigen erhebliche Variationen der Ergebnisse, je nach eingeschlossenen Studiendaten oder gewählten Analyseverfahren.
- Die Ergebnisse aus der Meta-Regressionsanalyse von Vermeulen et al. [44] sollten daher nicht vorbehaltlos in einer Risikoschätzung verwendet werden, insbesondere nicht im Niedrigexpositions-bereich gegenüber DME.

Korrespondenzadresse

PD Dr. P. Morfeld

Institut für Epidemiologie und Risikobewertung in der Arbeitswelt (IERA) der Evonik Industries AG, Rellinghauser Str. 1–11, 45128 Essen
peter.morfeld@evonik.com

Danksagung. Die Arbeiten von PD Dr. Peter Morfeld wurde durch Fördergelder der Europäischen Forschungsvereinigung für Umwelt und Gesundheit im Transportsektor (EUGT, www.eugt.org) unterstützt. EUGT e. V. hat sich zur Aufgabe gemacht, Aus- und Wechselwirkungen von Emissionen, Immissionen und Gesundheit zu untersuchen.

Einhaltung ethischer Richtlinien

Interessenkonflikt. P. Morfeld und M. Spallek weisen auf folgende Beziehungen hin: P. Morfeld ist Mitglied des wissenschaftlichen Beirats der Europäischen Forschungsvereinigung für Umwelt und Gesundheit im Transportsektor (EUGT e. V., www.eugt.org) und habilitiert für Epidemiologie an der Universität zu Köln. M. Spallek ist Geschäftsführer von EUGT e. V. und habilitiert für Arbeits- und Umweltmedizin an der Goethe-Universität Frankfurt/Main.

Dieser Beitrag beinhaltet keine Studien an Menschen oder Tieren.

Open Access Dieser Artikel unterliegt den Bedingungen der Creative Commons Attribution License. Dadurch sind die Nutzung, Verteilung, und Reproduktion erlaubt, sofern der/die Originalautor/en und die Quelle angegeben sind.

Literatur

- Al Khalaf MM, Thalib L, Doi SAR (2011) Combining heterogeneous studies using the random-effects model is a mistake and leads to inconclusive meta-analyses. *J Clin Epidemiol* 64:119–123
- Allison PD (2009) Fixed effects regression models. SAGE, Los Angeles
- Attfield MD, Schleiff PL, Lubin JH, Blair A, Stewart PA, Vermeulen R et al (2012) The diesel exhaust in miners study: a cohort mortality study with emphasis on lung cancer. *J Natl Cancer Inst* 104:869–883
- Benbrahim-Tallaa L, Baan RA, Grosse Y, Lauby-Secretan B, El Ghissassi F, Bouvard V et al (2012) Carcinogenicity of diesel-engine and gasoline-engine exhausts and some nitroarenes. *Lancet Oncol* 13:663–664
- Berlin JA, Longnecker MP, Greenland S (1993) Meta-analysis of epidemiologic dose-response data. *Epidemiology* 4:218–228
- Cameron AC, Trivedi PK (2005) Microeconometrics. Methods and applications. University Press, Cambridge
- Cameron AC, Trivedi PK (2010) Microeconometrics using Stata. Revised edition. Stata Press, College Station
- Checkoway H, Bofetta P, Mundt DJ, Mundt KA (2012) Critical review and synthesis of the epidemiologic evidence on formaldehyde exposure and risk of leukemia and other lymphohematopoietic malignancies. *Cancer Causes Control* 23(11):1747–1766
- Cherrie JW, Gorman Ng M, Shafrir A, van Tongeren M, Mistry R, Sobey M, Corden C, Rushton L, Hutchings S (2011) Health, socio-economic and environmental aspects of possible amendments to the EU Directive on the protection of workers from the risks related to exposure to carcinogens and mutagens at work. SHEcan Summary report. IOM Research Project: P937/99 May 2011. <http://ec.europa.eu/social/main.jsp?catId=716&langId=de&moreDocuments=yes>. Zugriffen: 21. Mai 2015
- Crump K (2014) Meta-analysis of lung cancer risk from exposure to diesel exhaust: study limitations. *Environ Health Perspect* 122:A230 (author's reply: 230–231). <http://dx.doi.org/10.1289/ehp.1408482>
- Crump K, Van Landingham C (2012) Evaluation of an exposure assessment used in epidemiological studies of diesel exhaust and lung cancer in underground mines. *Crit Rev Toxicol* 42:599–612
- Crump K, van Landingham C, Moolgavkar S, McClellan R (2014) Reanalysis of the DEMS nested case-control study of lung cancer and diesel exhaust: suitability for quantitative risk assessment. Health Effects Institute, Boston (Webinar September, 2014)
- Crump KS, van Landingham C, Moolgavkar SH, McClellan R (2015) Reanalysis of the DEMS nested case-control study of lung cancer and diesel exhaust: suitability for quantitative risk assessment. *Risk Anal*. doi:10.1111/risa.12371. <http://onlinelibrary.wiley.com/doi/10.1111/risa.12371/abstract>
- Diggle PJ, Heagerty PJ, Liang KY, Zeger SL (2002) Analysis of longitudinal data, 2. Aufl. Oxford University Press, Oxford
- Garshick E, Laden F, Hart JE, Davis ME, Eisen EA, Smith TJ (2012) Lung cancer and elemental carbon exposure in trucking industry workers. *Environ Health Perspect* 120:1301–1306
- Greenland S, Longnecker MP (1992) Methods for trend estimation from summarized dose-response data, with applications to meta-analysis. *Am J Epidemiol* 135:1301–1309
- Health Effects Institute (2015) Advanced Collaborative Emissions Study (ACES): lifetime cancer and non-cancer assessment in rats exposed to new-technology diesel exhaust. Research Report Number 184, January 2015. <http://pubs.healtheffects.org/view.php?id=430>. Zugriffen: 21. Mai 2015
- Hudson DJ (1966) Fitting segmented curves whose join points have to be estimated. *J Am Stat Assoc* 61:1097–1129
- Jones RH, Molitoris BA (1984) A statistical method for determining the breakpoint of two lines. *Anal Biochem* 141:287–290
- McClellan RO, Hesterberg TW, Wall JC (2012) Evaluation of carcinogenic hazard of diesel engine exhaust needs to consider revolutionary changes in diesel technology. *Regul Toxicol Pharmacol* 63:225–258
- Möhner M, Kersten N, Gellissen J. Diesel motor exhaust and lung cancer mortality: reanalysis of a cohort study in potash miners. *European Journal of Epidemiology* 2013;28(2):159–168.
- Moolgavkar S (2014) Diesel engine exhaust and lung cancer mortality – temporal factors in exposure and risk. Health Effects Institute, Boston (Webinar September, 2014)
- Moolgavkar SH, Chang ET, Luebeck G, Lau EC, Watson H, Crump K, McClellan RO (2015) Diesel engine exhaust and lung cancer mortality – time related factors in exposure and risk. *Risk Anal*. doi:10.1111/risa.12315. <http://onlinelibrary.wiley.com/doi/10.1111/risa.12315/abstract>
- Morfeld P (2012) Letter to Garshick et al. 2012: Cox model setup may lead to erroneous conclusions. *Environ Health Perspect* 120:A382 (author response A382–A383)
- Morfeld P (2012) Diesel exhaust in miners study: how to understand the findings? *J Occup Med Toxicol* 7(1):10. <http://www.occup-med.com/content/7/1/10>
- Morfeld P (2013) Letter to the editor – formaldehyde and leukemia: missing evidence! *Cancer Causes Control* 24:203–204 (Autorenantwort: 205)
- Morfeld P, Mundt KA, Taeger D, Guldner K, Steinig O, Miller BG (2013) Threshold value estimation for respirable quartz dust exposure and silicosis incidence among workers in the German porcelain industry. *J Occup Environ Med* 55:1027–1034
- Neter J, Wasserman W, Kutner MH (1985) Applied linear statistical models. Regression, analysis of variance and experimental designs. Richard Dr. Irwin, Homewood
- Neumeyer-Gromen A, Razum O, Kersten N, Seidler A, Zeeb H (2009) Diesel motor emissions and lung cancer mortality – results of the second follow-up of a cohort study in potash miners. *Int J Cancer* 124:1900–1906
- Orsini N, Bellocco R, Greenland S (2006) Generalized least squares for trend estimation of summarized dose-response data. *Stata J* 6:40–57
- Orsini N, Li R, Wolk A, Khudyakov P, Spiegelman D (2012) Meta-analysis for linear and nonlinear dose-response relations: examples, an evaluation of approximations, and software. *Am J Epidemiol* 175:66–73
- Park R (2014) Diesel engine emissions and risk assessment at NIOSH. Vortrag auf dem Workshop „Diesel Exhaust, Lung Cancer and Quantitative Risk Assessment“. Health Effects Institute, Boston (March 6, 2014)
- Rothman KJ, Greenland S, Lash TL (2008) Modern epidemiology, 3. Aufl. Lippincott Williams & Wilkins, Philadelphia
- Royston P, Altman DG, Sauerbrei W (2006) Dichotomizing continuous predictors in multiple regression: a bad idea. *Stat Med* 25:127–141
- Silverman DT, Attfield MD (2013) Letter-to-the-Editor RE: Diesel exhaust in miners study: how to understand the findings? by Peter Morfeld and author reply. <http://www.occup-med.com/content/7/1/10/comments>. Zugriffen: 21. Mai 2015
- Silverman DT, Samanic CM, Lubin JH, Blair AE, Stewart PA, Vermeulen R et al (2012) The diesel exhaust in miners study: a nested case-control study of lung cancer and diesel exhaust. *J Natl Cancer Inst* 104:855–868
- Stata Release 12 (2013) Statistical Software. StataCorp LP, College Station, Texas
- Steenland K, Silverman D, Hornung R (1990) Case-control study of lung cancer and truck driving in the Teamsters Union. *Am J Public Health* 80:670–674
- Steenland K, Deddens J, Stayner L (1998) Diesel exhaust and lung cancer in the trucking industry: exposure-response analyses and risk assessment. *Am J Ind Med* 34:220–228
- Stewart P, Coble JB, Vermeulen R, Schleiff P, Blair A, Lubin J, Attfield M, Silverman DT (2010) The diesel exhaust in miners study: I. Overview of the exposure assessment process. *Ann Occup Hyg* 54:728–746
- Sun Y, Bochmann F, Nold A, Mattenklott M (2014) Diesel exhaust exposure and the risk of lung cancer – a review of the epidemiological evidence. *Int J Environ Res Public Health* 11:1312–1340
- The Advisory Committee on Safety and Health at Work (2013) Supplementary opinion on the approach and content of an envisaged proposal by the Commission on the amendment of Directive 2004/37/EC on Carcinogens and Mutagens at the workplace. Supplementary opinion Doc. 727/13. Adopted on 30/05/2013. European Commission, DG Employment, Social Affairs and Inclusion Employment and Social Legislation, Social Dialogue, Health, Safety and Hygiene at Work. https://circabc.europa.eu/d/d/workspace/SpacesStore/85739c10-4b49-42ac-8982-aafc9a38f0eb/Doc%20%20727-13-%20EN%20Supplementary%20CMD%20Opinion_FINAL.pdf. Zugriffen: 21. Mai 2015
- Ulm K (1991) A statistical method for assessing a threshold in epidemiological studies. *Stat Med* 10:341–349
- Vermeulen R, Silverman DT, Garshick E, Vlaanderen J, Portengen L, Steenland K (2014) Exposure-response estimates for diesel engine exhaust and lung cancer mortality based on data from three occupational cohorts. *Environ Health Perspect* 122:172–177
- Zaebst D, Clapp D, Blade L, Marlow D, Steenland K, Hornung R, Scheutzle D, Butler J (1991) Quantitative determination of trucking industry workers exposures to diesel exhaust particles. *Am Ind Hygiene Assoc J* 52:529–541